Introduction to RADAR systems

Third Edition

Merrill I. Skolnik

McGRAW-HILL INTERNATIONAL EDITIONS
Electrical Engineering Series

chapter

6

Information from Radar Signals

6.1 INTRODUCTION

This chapter includes the *basic measurements* that can be made by a radar; the *theoretical accuracy* of radar measurements; the *ambiguity diagram* that graphically illustrates the characteristics of radar waveforms in the time (range) and frequency (radial velocity) domains; *pulse compression,* which is used to achieve high range-resolution without the need for high peak power; and target *recognition* methods whereby a radar distinguishes one type of echo signal from another.

A radar obtains information about a target by comparing the received echo signal with the signal that was transmitted. It was said in Chap. 5 that the presence of a target is announced when the echo signal is strong enough to cross the receiver detection threshold. Knowing that a target is present, however, is almost never sufficient in itself. More must be known to be useful; thus a radar must provide information about the target, as discussed next.

6.2 BASIC RADAR MEASUREMENTS

A radar can obtain a target's location in range and azimuth, and sometimes elevation. After several observations of a moving target over a period of time, the target trajectory, or track, can be obtained. Radar can also do more than simply characterize the target as a

"blob." In this section, the information available from a target first will be discussed as if the target were a point scatterer and then as a distributed scatterer. For purposes of this chapter, a *point scatterer,* or *point target,* is one with dimensions small compared to the size of the radar resolution cell in range, cross-range (angle), or both. The target's individual scattering features, therefore, are not resolved. A *distributed scatterer,* or *target* is one with dimensions large compared to the resolution-cell size, allowing the individual scatterers to be discerned. The resolution capabilities of a radar usually (but not always) determine whether a target is considered as a point scatterer (unresolved) or a distributed target (resolved). A *complex scatterer* is one that contains multiple scatterers. A complex scatterer can be either a point scatterer or a distributed scatter.

Measurements of a Point Target The basic radar measurements that can be made for a point target when only a single observation is made are range, radial velocity, direction (angle), and, in some special cases, tangential velocity.

Range It was said in Chap. 1 that the measurement of distance, or range, was obtained from the round-trip time T_R required for a radar signal to travel to the target and back. The range R is given by $cT_R/2$, where c = velocity of propagation. In many radar applications the target's range is the most significant measurement that is made. No other sensor has been able to compete with radar for determining range to a distant target, especially in accuracy, ability to make a measurement over very long or very short distances, and under adverse weather conditions. A long-range air-surveillance radar might measure range to an accuracy of many tens of meters, but accuracies of a few centimeters are possible with precision systems. In the most precise systems, the accuracy of a range measurement is limited only by the accuracy with which the velocity of propagation is known. The spectral bandwidth occupied by the radar signal is the fundamental resource required for accurate range measurement. The greater the bandwidth, the more accurate can be the range measurement.

Angle Measurement Almost all radars utilize directive antennas with relatively narrow beamwidths. A directive antenna not only provides the large transmitting gain and large receiving aperture needed for detecting weak echo signals, but its narrow beamwidth allows the target's direction to be determined accurately. It can do this by noting the direction the antenna points when its received echo signal is a maximum. A typical microwave radar might have a beamwidth of one or a few degrees. The narrowest beamwidths of operational radars have been about 0.3°. This is not an absolute limit; but the narrower the beamwidth, the greater the mechanical and electrical tolerances that are required of the antenna.

Angular accuracy can be much better than the antenna beamwidth, as described later in this chapter. Angle accuracy depends on the electrical size of the antenna (size as measured in wavelengths). With signal-to-noise ratios typical of those required for reliable detection, the angular location of a target can be determined to about 1/10th of a beamwidth. The best precision monopulse tracking radars used for range instrumentation can determine angle to about 0.1 mrad rms (0.006°) if the signal-to-noise ratio is large enough and if the proper efforts are taken to minimize errors.

Radial Velocity Measurement of the radial component of velocity in many radars is obtained from the rate of change of range. This is known as the *range rate.* The classical method for finding the radial velocity is based on $v_r = (R_2 - R_1)/(T_2 - T_1)$. It is found from the range R_1 measured at time T_1 and the range R_2 at time T_2. However, this method of finding range rate (or any other derivative of a location measurement) is not considered here as a *basic* radar measurement even though it may be widely used. Instead, the doppler frequency shift is the basic method for obtaining radial velocity. It can be made on the basis of a single observation. Using the classical expression [Eq. (3.3)] for the doppler frequency shift, f_d, the radial velocity v_r is given as

$$v_r = \lambda f_d/2 \qquad [6.1]$$

where λ = wavelength. It can be shown from the theoretical accuracy expressions presented later in this chapter that the radial velocity accuracy derived from the doppler frequency shift can be much better than that found from the range rate, assuming the time between the two range measurements in the range-rate method is the same as the time duration of the doppler frequency measurement. (See Problem 6.5.)

It will be seen later that the accuracy of the doppler-frequency measurement depends on the time duration over which it is made. The longer the time, the better the frequency accuracy. Because of the relationship between radial velocity and wavelength in Eq. (6.1), the shorter the wavelength, the shorter can be the observation time to achieve a required velocity accuracy. (The shorter the wavelength, the higher the frequency.) Or, the shorter the wavelength, the better will be the velocity accuracy for a given observation time.

In spite of its good accuracy, the doppler frequency shift is not used as often as is the range-rate for obtaining the radial velocity since it can result in ambiguities in range and/or doppler when employed with a short or a medium pulse-duration radar.

Tangential (Cross-Range) Velocity Just as the temporal doppler frequency shift can provide the radial velocity, there exists in the spatial (angle) domain an analogous *spatial doppler-frequency shift* from which the tangential velocity can be determined.[1] (If the radial velocity is $v_r = v \cos \theta$, the tangential velocity is $v_t = v \sin \theta$, where v is the target's speed and θ is the angle between the target's velocity vector and the radar line of sight.) The angle-rate times the range is equal to the tangential velocity. Together, the tangential velocity and the radial velocity can give the magnitude of the target's speed v and its direction θ. The measurement of tangential velocity has not been of practical interest in radar since it requires a long-baseline antenna system.

Measurements of a Distributed Target With sufficient resolution in the appropriate dimension, the size and shape of a distributed target can be ascertained. It should be recalled that resolution and accuracy are not the same. Range *resolution* requires that the entire bandwidth be occupied continuously without gaps in the signal frequency spectrum. Range *accuracy,* however, only requires, as a minimum, that there be adequate spectral energy at the two ends of the spectral bandwidth. The spectral bandwidth need not be fully occupied. This assumes that there is only one scattering object present. Resolution requires a filled spectrum; accuracy can be achieved with a thinned, or sparse, spectrum. A similar description applies in the temporal (time) domain for frequency measurement,

and in the spatial (antenna) domain for angle measurements. Generally, good resolution will provide good accuracy; but the reverse is not necessarily true since accurate measurements can be made with waveforms that do not provide good resolution.

Radial Profile The target's profile (and size) in the range dimension can be obtained when the radar's range resolution cell is smaller in size than the target's dimensions (i.e., when the scattering centers of the target can be resolved). To obtain a radial profile of a target it is required that $c\tau/2 << D$, where D = the target's radial component and τ = pulse width. Good resolution in the range dimension requires large spectral bandwidth. The radial profile of a target sometimes can be employed to obtain limited "recognition" of one type of target from another.

Tangential (Cross-Range) Profile With sufficient resolution in the angle dimension, the tangential (cross-range) profile of a distributed target can be determined. This can provide the angular size of the target and the location in angle of the scattering centers. If the range is known, the location of scatterers in the tangential dimension can be determined since the cross-range (tangential) dimension is equal to the product of range and angle (the latter in radians). Resolution in cross-range based on conventional angle measurements is generally not as good as the resolution that can be obtained in the range dimension. Synthetic aperture radar (SAR) and inverse synthetic aperture radar (ISAR), however, can provide excellent cross-range resolution without the need for large antennas. (In SAR and ISAR, the equivalent of resolution in angle may be thought of as being obtained because of resolution in the doppler frequency domain.)

Size and Shape When the tangential profile is obtained at each range resolution cell, the target image (size and shape) is formed. Imaging radars, such as SAR, ISAR, and SLAR (side-looking airborne radar), have sufficient resolution in both range and cross-range to resolve the major scatterers of a distributed target. (SLAR achieves its tangential resolution by use of a narrow beam antenna directed perpendicular to the flight direction of the aircraft carrying the radar.)

Symmetry The response of a target to changes in the polarization of the radar signal can provide a measure of the symmetry of the target. (The polarization of a radar signal is determined by the orientation of the electric field.) If a sphere (a perfectly symmetrical target) were directly viewed by a radar with a rotating linearly polarized signal, there would be no change of the echo signal when the polarization is changed. On the other hand, if the same rotating polarization radar were to view a long narrow rod, the echo would be maximum when the electric field (polarization) is parallel to the rod and minimum when it is perpendicular to the rod. By observing the variation of the amplitude of the echo signal as a function of polarization, the orientation and shape of the rod can be determined. Measurement of target symmetry using polarization is not widely used in radar; however, it is the basis for detection of aircraft (an asymmetrical target) in the presence of rain (symmetrical target) when using circular polarization (defined as the electric field rotating at the RF frequency, Sec. 7.8).

Change of Radial and Tangential Profiles Here it is assumed that the pulse is long enough so that the individual scatterers of a complex target are not resolved. If the individual scatterers of this complex target change their relative locations in range (radial profile), the echo signal will experience a change in amplitude due to constructive and destructive interference among the echo signals from the individual scatterers. Changes in the target echo amplitude, therefore, indicate there are changes in the relative locations of the individual scatterers of the complex target.

Examples of target effects that might be recognized by the amplitude modulation of the echo signal include changes in target aspect, propeller modulation, jet engine modulation, and the time-varying separation of two closely spaced, unresolved targets (such as two aircraft or an aircraft and a missile).

Surface Characteristics The dielectric constant of a target's surface material and the roughness of its surface can, in principle, be found from radar measurements. Surface roughness may be determined by varying the radar frequency and noting where the scattering changes from specular (a smooth surface) to diffuse (rough surface). This boundary depends on the size of the surface roughness relative to the radar wavelength. Surface roughness, such as the height of ocean waves (the sea state), can be found from a direct measurement with a high range-resolution radar, as has been done from space with a precision high-resolution altimeter.

The dielectric constant of the scattering surface can be found if the reflection coefficient can be measured and if the shape and roughness of the surface are known. This is practical under laboratory conditions, but difficult to apply with radar. Radar cross section measurements over a wide range of frequencies, however, were used to estimate the dielectric properties of the moon's surface (before astronauts landed on the moon and brought back rocks for laboratory analysis).[2,3]

The surface roughness and the dielectric constant are of interest for remote sensing with radar, especially from space. The former might indicate the sea state over the oceans of the world; and the latter, if it were practical, might be used to determine soil moisture, which is of interest for agriculture and hydrology. Although the radar determination of surface characteristics might be desirable, it has proven to be difficult to achieve except under limited circumstances.

6.3 THEORETICAL ACCURACY OF RADAR MEASUREMENTS

Noise is the fundamental limitation to accurate radar measurements. The theoretical aspects of the extraction of information from radar signals have benefited greatly from the theory of *statistical parameter estimation* just as the theory of detection has benefited from the statistical theory *of hypothesis testing.*[4] In this section, expressions for the theoretical accuracies of radar measurements will be presented. It is assumed that the signal-to-noise ratio is large. This is usually the case since it was found in Sec. 2.5 that large signal-to-noise ratios are required for detection of a signal. Detection must occur before

meaningful information can be extracted about a target echo. It is further assumed that the measurement error associated with a particular parameter is independent of the errors in any of the other parameters, that accuracy is limited only by receiver noise, and that all bias errors are accounted for separately. The measure of error is the root mean square (rms) of the difference between the measured (estimated) value and the true value.

The expressions given in this chapter for the theoretical rms error δM of a radar measurement M have the following form:

$$\delta M = \frac{kM}{\sqrt{2E/N_0}} \qquad [6.2]$$

where k is a constant whose value is in the vicinity of unity, E is the received signal energy, and N_0 is the noise power per unit bandwidth. The following will be shown later in this section:

- For a time-delay (range) measurement, k depends of the shape of the frequency spectrum $S(f)$, and M is the rise time of the pulse (inversely proportional to bandwidth).
- For a measurement of radial velocity based on the doppler frequency, k depends on the shape of the time waveform $s(t)$, and M is the spectral resolution (inversely related to the time duration of the signal).
- For an angle measurement, k depends on the shape of the aperture illumination $A(x)$, and M is the beamwidth.

Theoretical radar accuracies may be derived by a variety of methods such as (1) simple geometrical relationships among the signal, noise, and the parameter to be measured;[5] (2) the likelihood ratio;[6] (3) the method of inverse probability;[7] (4) a suitably selected gating function preceded by a matched filter;[8] and (5) minimization of the mean square error.[9]

The simple method (no. 1) for finding the rms error in the measurement of time delay when the waveform is a rectangular pulse will be illustrated next. This derivation takes some liberties, but it has the advantage of being easy to understand. Fortunately, the simple method and the more involved methods give similar answers for the rectangular pulse.

Time-Delay (Range) Accuracy—Simplified Method The measurement of range R is the measurement of the round-trip time delay T_R for the radar signal (waveform) to travel out to the target and back. The rms error in range is $\delta R = (c/2)\ \delta T_R$, where c is the velocity of propagation, and δT_R is the rms error in time delay. The range, or time-delay, measurement to be described here is based on locating the leading edge of the video pulse, Fig. 6.1. The video pulse uncorrupted by noise is shown by the solid curve. Its shape is not perfectly rectangular, but has finite rise and fall times. (Zero rise or fall times require infinite bandwidth.) The effect of noise added to the pulse is to shift the time of threshold crossing as shown by the dashed curve. Since large signal-to-noise ratio is assumed, the slope of the leading edge of the noise-free pulse (solid curve) can be equated to the slope of the leading edge of the pulse with noise added (dashed curve). The slope of the leading edge of a pulse of amplitude A at the output of a video filter is A/t_r, where t_r is the rise time. From Fig. 6.1, the slope of the signal plus noise (dashed curve) can be

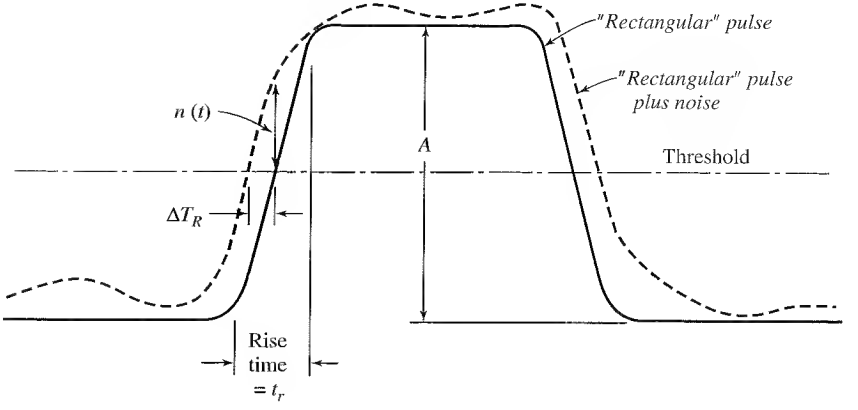


Figure 6.1 Measurement of time delay using the leading edge of the video pulse. Solid curve is the echo pulse uncorrupted by noise. Dashed curve represents signal plus noise.

written as $n(t)/\Delta T_R$, where $n(t)$ is the noise voltage at the threshold crossing of the pulse, and ΔT_R is the error in the time-delay measurement. Equating the slope of the leading edge of the pulse without noise to the slope of the pulse with noise gives

$$A/t_r = n(t)/\Delta T_R \qquad [6.3]$$

which leads to

$$[\overline{(\Delta T_R)^2}]^{1/2} = \delta T_R = \frac{t_r}{(A^2/\overline{n^2})^{1/2}} = \frac{t_r}{(2S/N)^{1/2}} \qquad [6.4]$$

where $A^2/\overline{n^2}$ is the video signal-to-noise power ratio. The last part of Eq. (6.4) follows from the fact that the signal-to-noise power ratio for a rectangular video pulse is equal to $2S/N$, where S/N is the signal-to-noise power ratio of a sinewave pulse in the IF portion of the receiver. This assumes a linear detector and a large signal-to-noise ratio. Equation (6.4) indicates that accurate measurements of time delay require video pulses with short rise times and large amplitudes. The width of the pulse does not enter explicitly in this expression.

If the rise time of the video pulse is limited by the spectral bandwidth B of the rectangular-shaped IF filter, then $t_r \approx 1/B$. Letting $S = E/\tau$ and $N = N_0B$, the error in the time delay can be written

$$\delta T_R = \left(\frac{\tau}{2BE/N_0}\right)^{1/2} \qquad [6.5]$$

where τ = pulse width, B = spectral bandwidth of the rectangular filter, E = signal energy, and N_0 = noise power per unit bandwidth. If a similar measurement of time delay is made at the trailing edge of the video pulse, and if the noise at the trailing edge is independent of the noise at the leading edge, then the two measurements can be averaged to obtain an improvement in the time-delay accuracy of $\sqrt{2}$, which gives

$$\delta T_R = \left(\frac{\tau}{4BE/N_0}\right)^{1/2} \qquad [6.6]$$

The estimate of the time delay obtained by determining when either the leading edge or trailing edge of the pulse crosses a threshold, as in Fig. 6.1, will depend on the value of the threshold relative to the peak value of the pulse. The choice of threshold level is important, therefore, for consistency in accurate measurement when only the leading edge is used. It has been suggested that the bias error can be avoided by one of several methods.[10–12] One method is to use an adaptive threshold in which the level of the threshold is always a fixed fraction of the pulse amplitude. If the average of the two time delays found from both the leading and trailing edges of the pulse is used, there is no theoretical bias with change of amplitude if the pulse shape is symmetrical.

Time-Delay Accuracy and Effective Bandwidth There are several methods[6–9] based on the likelihood ratio, inverse probability, and other statistical analyses, which all lead to the following expression for the rms error in the measurement of the time delay

$$\delta T_R = \frac{1}{\beta\,(2E/N_0)^{1/2}} \qquad [6.7]$$

where E is the signal energy, N_0 is the noise power per unit bandwidth, and β is called the *effective bandwidth* and is defined as

$$\beta^2 = \frac{\int_{-\infty}^{\infty} (2\pi f)^2 |S(f)|^2\, df}{\int_{-\infty}^{\infty} |S(f)|^2\, df} = \frac{1}{E} \int_{-\infty}^{\infty} (2\pi f)^2\, |S(f)|^2\, df \qquad [6.8]$$

It has also been called the *rms bandwidth.* The effective bandwidth β is such that $(\beta/2\pi)^2$ is the normalized second moment of $|S(f)|^2$ about its mean. Equation (6.8) assumes that the mean value of $S(f)$ is at $f = 0$, where $S(f)$ is the video spectrum with negative as well as positive frequencies. The effective bandwidth β is different from other bandwidths encountered in electronic engineering. It is not related to either the half-power bandwidth or the noise bandwidth. The more the spectral energy is concentrated at the two ends of the band, the larger will be β and the more accurate will be the measurement of time delay.

The first edition (1962) of this text discussed three different methods for deriving Eq. (6.8) based on statistical concepts.[6–8] The second edition (1980) described one method to derive this equation.[8] These methods are not included in the current edition since there now appears to be less interest in the mathematical aspects of the subject. The application of Eq. (6.8), however, is important and will be discussed next.

Rectangular Pulse When the spectrum $S(f)$ of a perfectly rectangular pulse—one with zero rise time and zero fall time—is inserted in Eq. (6.8) for the effective bandwidth, the result is obtained that $\beta = \infty$. This means that $\delta T_R = 0$; hence, the measurement of the time delay can be made with zero error. It may seem strange, but it is correct for the perfectly rectangular pulse that was assumed. An infinite bandwidth implies zero rise time (infinite slope) so noise does not displace the threshold crossing in time (as it does in Fig. 6.1 for a finite rise time) and there will be no error in the delay. A perfectly rectangular

pulse, however, requires infinite bandwidth, which is not possible. Thus the bandwidth of a practical "rectangular" pulse must be finite, there will be finite rise and fall times, and the rms time-delay error will not be zero.

To obtain the effective bandwidth β for a finite-bandwidth pulse, it will be assumed that the spectrum of a perfectly rectangular IF pulse of width τ_r is limited to a finite spectral bandwidth B_s. For the time-delay measurement that uses the envelope of the IF pulse, this is equivalent to a video spectrum $S(f) = (\sin \pi f\tau_r)/\pi f\tau_r$ that is limited to a spectral width $\pm B_s/2$, as in Fig. 6.2. (The video spectrum is shown here with both negative and positive frequencies, as is required in Fourier analysis.) Although the analysis considered here is based on the video pulse of width τ_r and a low-pass filter of video bandwidth $B_s/2$, the result is the same as taking the envelope of an IF pulse of width τ_r and an IF bandpass filter of bandwidth B_s. The value of β^2 for this case is found by setting the limits of the integration in Eq. (6.8) from $-B_s/2$ to $+B_s/2$ *instead* of from $-\infty$ to $+\infty$, which is then

$$\beta^2 = \frac{(2\pi)^2 \int_{-B_s/2}^{B_s/2} f^2(\sin^2 \pi f\tau_r)/\pi^2 f^2 \, df}{\int_{-B_s/2}^{B_s/2} (\sin^2 \pi f\tau_r)/\pi^2 f^2 \, df} = \frac{1}{\tau_r^2} \frac{\pi B_s\tau_r - \sin \pi B_s\tau_r}{\text{Si}(\pi B_s\tau_r) + (\cos \pi B_s\tau_r - 1)/\pi B_s\tau_r} \quad [6.9]$$

where $\text{Si}(x)$ is the *sine integral* function defined by $\int_0^x (\sin u)/u \, du$. For large $B_s\tau_r$ in Eq. (6.9), the product $\beta^2\tau_r^2 \to 2B_s\tau_r$, or

$$\beta^2 \approx \frac{2B_s}{\tau_r} \approx \frac{2}{\tau_r t_r} \qquad \text{for large } B_s\tau_r \quad [6.10]$$

It was assumed in Eq. (6.10) that the rise time t_r of the pulse is approximately the inverse of the spectral bandwidth B_s (the total width of the spectrum, not the half-power bandwidth).

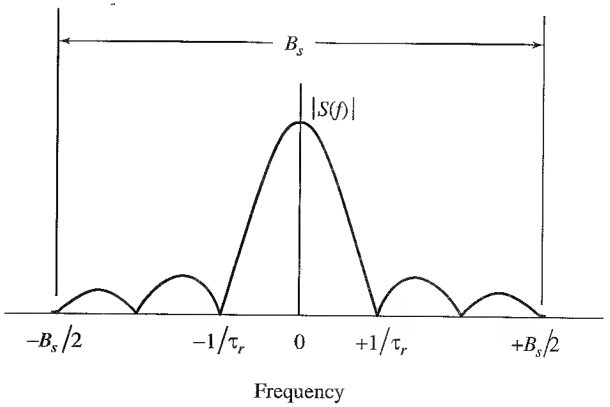


Figure 6.2 Spectrum $[(\sin \pi f\tau_r)/\pi f\tau_r]$ of a rectangular pulse of width τ_r shown limited to a spectral extent of $\pm B_s/2$.

Substituting this expression for β^2 into Eq. (6.7) gives the rms error in time delay as

$$\delta T_R = \left(\frac{\tau_r}{4B_s E/N_0}\right)^{1/2} = \left(\frac{t_r \tau_r}{4E/N_0}\right)^{1/2} \qquad \text{rectangular pulse, } B_s\tau_r >> 1 \qquad [6.11]$$

This applies for large $B_s\tau_r$, or when the rise time is small compared to the pulse width. Note that Eq. (6.11), derived in a totally different manner, is the same as Eq. (6.6) except for denoting the spectral bandwidth in Eq. (6.11) as B_s instead of B and the pulse width as τ_r instead of τ.

The value of $\beta = (2B_s/\tau_r)^{1/2}$ for a long rectangular pulse of fixed bandwidth B_s (or fixed rise time) decreases with increasing pulse width τ_r. Thus if the total energy remains the same, the time-delay accuracy decreases (becomes worse) with increasing pulse width even though the rise time remains the same.

It is not often that $B_s\tau_r >> 1$ in radar applications. Next, the more usual case is examined, where the product of half-power bandwidth and pulse width is approximately unity.

Quasi-Rectangular Pulse As before, we start with a perfectly rectangular pulse of width τ_r. The spectral extent B_s is assumed in Fig. 6.2 to be limited to the main portion of the rectangular video pulse spectrum that lies between the first nulls at $-1/\tau_r$ and $+1/\tau_r$ on either side of the spectrum peak at $f = 0$. (As mentioned, integration over negative as well as positive frequencies has to be considered in Fourier analysis). Thus the IF spectral bandwidth extent is $B_s = 2/\tau_r$. The half-power bandwidth is $B \approx B_s/2$; or $B \approx 1/\tau_r$. (The product of the half-power bandwidth B and the width τ_r of a rectangular pulse is actually equal to 0.886; but, for convenience, it is usually "rounded off" to unity. This is more like the usual case in radar where $B\tau \approx 1$.) The solid curve of Fig. 6.3 shows the pulse shape that

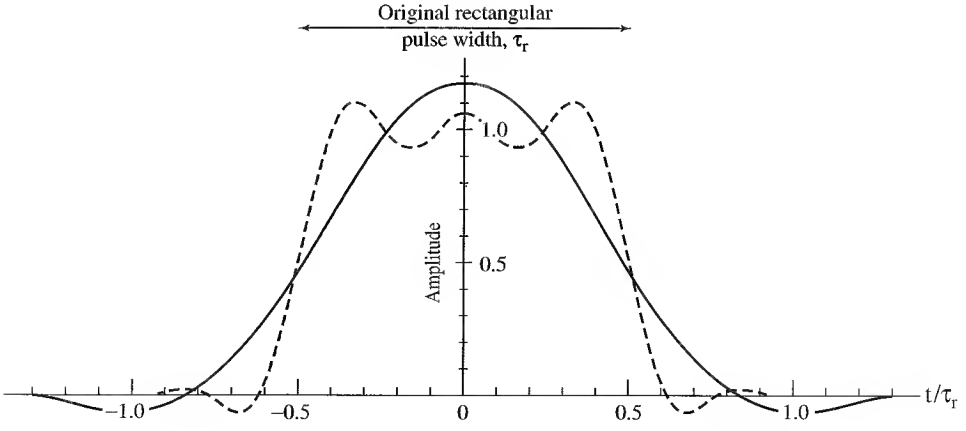


Figure 6.3 Solid curve is the theoretical waveform out of a low-pass rectangular filter of bandwidth $B_v = B_s/2 = 1/\tau_r$ when the input is a rectangular pulse of width τ_r. It is the same as the envelope out of an IF filter of bandwidth $B_s = 2/\tau_r$ when the input is a rectangular pulse of sinewave of width τ_r. In the text this is called a *quasi-rectangular pulse*. The dashed curve applies for $B_s = 6/\tau_r$.

emerges from the video low-pass filter of bandwidth $B_s/2$, which corresponds to an IF filter with bandwidth B_s. Although Fig. 6.3 doesn't resemble a perfectly rectangular pulse, it will be called the *quasi-rectangular pulse.* It is not unusual for a pulse radar to radiate a similar-looking waveform when it is thought that a "rectangular" pulse is being generated. (A rounded pulse is often used in radar since it produces less out-of-band interference to other users of the electromagnetic spectrum.) Substituting $B_s\tau_r = 2$ into Eq. (6.9), we get that $\beta\tau_r$ is very close to 2.1; which, when inserted into Eq. (6.7), gives the rms time-delay error as

$$\delta T_R = \frac{\tau_r}{2.1(2E/N_0)^{1/2}} \qquad \text{quasi-rectangular pulse, } B\tau_r \approx 1 \qquad \textbf{[6.12]}$$

This expression for the rms time-delay error is considered here to be representative of a conventional pulse typically used in radar. Since the rectangular IF filter is of bandwidth $B_s = 2/\tau_r$, then $\beta = 1.05B_s$. In terms of the half-power bandwidth B of the IF spectrum, $\beta = 2.4B$ when $B\tau_r = 0.886$, or $\beta \approx 2.1B$ when $B\tau \approx 1$. The dashed curve in Fig. 6.3 for $B_s\tau_r = 6$ is shown to indicate the shorter rise time of the pulse when the bandwidth B_s is much wider than the quasi-rectangular pulse.

The symbol τ_r was used for the pulse width in the above analysis instead of τ to indicate it was the width of the perfectly rectangular pulse before it is passed through a rectangular IF filter of bandwidth B_s (or video bandwidth $B_s/2$). If τ denotes the half-power pulse width that emerges from a filter of bandwidth $B_s = 2/\tau_r$, then $\tau = 0.625\tau_r$. The effective bandwidth in this case is $\beta = 1.3/\tau$.

Trapezoidal Pulse A rectangular pulse is sometimes approximated by a trapezoidal pulse with finite rise and fall times. If the width across the top of the trapezoid is τ_p and if the rise and fall times are t_r, then the rms error in time-delay measurement is

$$\delta T_R = \left(\frac{3\tau_p t_r + 2t_r^2}{12E/N_0}\right)^{1/2} \qquad \text{trapezoidal pulse} \qquad \textbf{[6.13]}$$

When the rise time t_r is small compared to the width τ_p of the top of the trapezoid, Eq. (6.13) becomes

$$\delta T_R = \left(\frac{\tau_p t_r}{4E/N_0}\right)^{1/2} \qquad t_r << \tau_p \qquad \textbf{[6.14]}$$

which is the same as Eq. (6.11) for the rectangular pulse with short rise time. Note that the larger the width τ_p, the poorer is the accuracy.

Triangular Pulse When $\tau_p = 0$ in Eq. (6.13), the pulse is triangular and its base = $2t_r = \tau_B$. The rms error becomes

$$\delta T_R = \frac{\tau_B}{12^{1/2}(2E/N_0)^{1/2}} \qquad \text{triangular pulse} \qquad \textbf{[6.15]}$$

In terms of the half-power bandwidth B of the $(\sin^2 x)/x^2$ spectrum of a triangular pulse (where $x = \pi f\tau_B/2$), the effective bandwidth $\beta = 2.72\ B$.

With a triangular or rounded pulse (such as the quasi-rectangular, gaussian, or $(\sin x)/x$ time waveforms) the time delay might be found by differentiating the pulse waveform and selecting the time at which $s'(t)$ goes through zero (which is the point at which the amplitude of $s(t)$ is a max).

Gaussian Pulse Consider a pulse described by the gaussian function

$$s(t) = \exp\left(-\frac{1.38t^2}{\tau^2}\right) \qquad [6.16]$$

where τ in this case is the half-power pulse width. Its rms time-delay error is

$$\delta T_R = \frac{\tau}{1.18(2E/N_0)^{1/2}} = \frac{1.18}{\pi B(2E/N_0)^{1/2}} \qquad \text{gaussian pulse} \qquad [6.17]$$

where B is the half-power bandwidth of the gaussian-pulse spectrum.

Cosine Pulse The positive half of one cycle of a cosine also can represent practical radar pulses. Its time waveform is $\cos(\pi t/\tau_B)$, where τ_B is the width of the base. Its effective bandwidth is $\beta = \pi/\tau_B = 2.64B$, where B = half-power bandwidth. The parabolic pulse, given by $s(t) = 1 - 4t^2/\tau_B^2$, is similar in shape and has $\beta = 3.16/\tau_B$.

Pulse with Uniform Spectrum The effective bandwidth of a waveform with a uniform spectrum of width B_u is $\beta = \pi B_u/\sqrt{3}$. The time waveform is of the form $(\sin \pi B_u t)/\pi B_u t$. The rms time-delay error is

$$\delta T_R = \frac{\sqrt{3}}{\pi B_u (2E/N_0)^{1/2}} \qquad \frac{\sin \pi B_u t}{\pi B_u t} \text{ waveform} \qquad [6.18]$$

The above applies to a linear-FM pulse-compression waveform with large time-bandwidth product.

Minimum Error Examination of the integral in the numerator of the expression for β^2 in Eq. (6.8) indicates that a large value of β requires that the spectral energy be concentrated at the extremes of the spectral bandwidth [note that the origin is at the mean value of $S(f)$]. Consider, therefore, a spectrum with its entire energy concentrated in delta functions at both ends of the video spectrum, such as $S(f) = \delta(f + B_s/2) + \delta(f - B_s/2)$. Substitution of this spectrum into Eq. (6.8) yields $\beta = \pi B_s$. This is the largest value of β that can be obtained for a spectrum occupying a bandwidth B_s. The time waveform corresponds to two sinewaves, each of infinite duration, separated in frequency by B_s. (The time duration of the waveform, of course, cannot be infinite. The result is approximately the same, however, if the time duration τ is such that $1/\tau << B_s$.) This waveform, however, results in range ambiguities that have to be resolved. It does not work when multiple unresolved targets are present.

Two CW waveforms separated in frequency by B have been considered in the past for measurement of range by determining the phase difference between them.[13,14] The phase difference $\Delta\phi$ between the two frequencies separated by B gives the time delay as

$$T_R = \frac{\Delta\phi}{2\pi B} \qquad [6.19]$$

The error in time delay, δT_R, based on the theoretical error in phase $\delta\phi$ [which is $\delta\phi = (2S/N)^{-1/2}$] from each of the two sine waves (with each having one half the total energy) results in the same value of $\beta = \pi B$ as found above by substituting the two delta-function spectrum into Eq. (6.8).

The two-frequency CW waveform might have a higher effective bandwidth than the others (and better time-delay accuracy), but it is inconvenient to use because of ambiguities. In practical embodiments of this method of range measurement, four or five frequencies have been used to resolve the ambiguities and achieve good accuracy.[15]

Other Considerations In the expression for time-delay error as given by Eqs. (6.7) and (6.8), the energy E appears as a normalizing factor in the effective bandwidth β. It also appears in the signal-to-noise energy ratio ($2E/N_0$). The energy cancels, and the rms time-delay error can be written as

$$\delta T_R = \left(\frac{N_0}{2\int_{-\infty}^{\infty} (2\pi f)^2 |S(f)|^2 \, df} \right)^{1/2} \qquad [6.20]$$

The integral in the denominator also can be expressed as

$$\beta^2 E = \int_{-\infty}^{\infty} [s'(t)]^2 \, dt = -\int_{-\infty}^{\infty} s''(t) s(t) \, dt \qquad [6.21]$$

The energy E does not appear in the right-hand side of this equation for $\beta^2 E$ since $\beta^2 \sim 1/E$. When calculating β, the expressions of Eq. (6.21) sometimes are easier to compute than the integral of Eq. (6.8). Equation (6.21) implies that accurate time delay is obtained with waveforms having large first derivatives over the time duration of the signal.

Table 6.1 lists the effective bandwidths β for the various waveforms considered here. There is not much difference in β among the various waveforms in this table. One might, therefore, not be too concerned about which waveform ought to be chosen based on time-delay accuracy alone. The triangular waveform has a large theoretical accuracy, but the discontinuity of the slope at the middle of the pulse presents practical problems in its generation. The values of β for the "rounded" pulses (gaussian, cosine, and quasi-rectangular) are not much lower than the triangular, and they are better representations of the shape of practical radar pulses. Radar pulses are almost always band-limited and are, therefore, rounded rather than appear as the perfectly rectangular pulses as sometimes seen in textbooks. The cosine seems a good compromise choice as typical of these rounded waveforms. The uniform spectrum, such as that of the linear-FM pulse-compression waveform, is about $2/3d$ less accurate than the rounded pulse if its spectral extent B equals the half-power bandwidth of the rounded pulse, but this is not much of a difference.

Accuracy of Frequency and Radial Velocity The measurement of frequency in radar is that of the doppler frequency shift. As mentioned previously, the radial component v_r of velocity can be found from the doppler frequency shift $f_d = 2v_r/\lambda$, where λ is the radar

Table 6.1 Effective Bandwidth β for Various Waveforms

Description	Time Waveform $s(t)$	Effective Bandwidth β
Gaussian pulse	$\exp(-1.38t^2/\tau^2)$	2.66 B, or $1.18/\tau$
Cosine pulse	$\cos(\pi t/\tau_B)$	2.64 B, or $3.14/\tau_B$
Triangular pulse	$(2t/\tau_B)+1 \quad -\tau_B/2 < t < 0$ $-(2t/\tau_B)+1 \quad 0 < t < \tau_B/2$	2.72 B, or $3.46/\tau_B$
Quasi-rectangular* $B_s\tau_r = 2$; $B\tau_r = 0.886$	$\mathrm{Si}[\pi(B_s t+1)] - \mathrm{Si}[\pi(B_s t-1)]$	2.38 B, or $2.1/\tau_r$, or $1.3/\tau$
Uniform spectrum of width B_u	$(\sin \pi B_u t)/\pi B_u t$	1.8 B_u, or $2.04/\tau$
Band-limited, rectangular pulse; $B_s\tau_r >> 1$	$\mathrm{Si}[\pi B_s(t+\tau_r/2)] - \mathrm{Si}[\pi B_s(t-\tau_r/2)]$	$1.4\sqrt{B_s/\tau_r}$
2 CW sinewaves separated by B_s	video spectrum: $\delta(f+B_s/2)+\delta(f-B_s/2)$	3.14 B_s

B_s = spectral extent, B = half-power bandwidth, τ = half-power pulsewidth, τ_r = width of original rectangular pulse, τ_B = extent of pulse (at its base), Si [X] = sine integral function of X.

*Considered to be typical of many pulse waveforms.

wavelength. The rms error in radial velocity is $\delta v_r = (\lambda/2)\,\delta f_d$, where δf_d is the rms error in the doppler frequency.

Using the method of inverse probability, Roger Manasse[16] showed that the rms error in the measurement of frequency is

$$\delta f = \frac{1}{\alpha\,(2E/N_0)^{1/2}} \qquad [6.22]$$

where

$$\alpha^2 = \frac{\int_{-\infty}^{\infty}(2\pi t)^2 s^2(t)\,dt}{\int_{-\infty}^{\infty} s^2(t)\,dt} \qquad [6.23]$$

and $s(t)$ is the input signal as a function of time. Note the similarity of this expression for δf and that of δT_R in Eq. (6.7), as well as the similarity in the expressions for α and β. The parameter α is the *effective time duration* of the signal, and $(\alpha/2\pi)^2$ is the normalized second moment of $s^2(t)$ about the mean epoch, taken to be $t = 0$. [If the mean is not zero, but is some other value, t_0, the integrand in the numerator of Eq. (6.23) would be $(2\pi)^2(t-t_0)^2 s^2(t)$.]

Rectangular Pulse The value of α^2 for a perfectly rectangular pulse of width τ is found to be $\pi^2\tau^2/3$; thus the rms frequency error is

$$\delta f = \frac{\sqrt{3}}{\pi\tau(2E/N_0)^{1/2}} \qquad \text{rectangular pulse} \qquad \textbf{[6.24]}$$

The longer the pulse, the more accurate is the frequency measurement. This expression can also be applied to a frequency measurement made by a CW radar since the observation time over which the CW measurement is made is equivalent to the pulse duration τ.

Quasi-Rectangular Pulse For the bandwidth-limited rectangular pulse of Fig. (6.2), the value of α^2 is

$$\alpha^2 = \frac{\pi^2\tau_r^2}{3} \frac{\text{Si}(\pi B_s\tau_r) + \dfrac{\cos(\pi B_s\tau_r) - 3}{\pi B_s\tau_r} - \dfrac{2\sin\pi B_s\tau_r}{(\pi B_s\tau_r)^2} - \dfrac{8[\cos(\pi B_s\tau_r) - 1]}{(\pi B_s\tau_r)^3}}{\text{Si}(\pi B_s\tau_r) + [\cos(\pi B_s\tau_r) - 1]/\pi B_s\tau_r} \qquad \textbf{[6.25]}$$

where τ_r is the width of a rectangular pulse that is passed through a rectangular filter of bandwidth B_s, and Si (x) is the sine integral function of x. In the limit as $B_s\tau_r \to \infty$, α^2 approaches $\pi^2\tau_r^2/3$, which is the same as that obtained for the perfectly rectangular pulse.

In the discussion of the rms error in time delay given previously, a *quasi-rectangular pulse* with $B_s\tau_r = 2$ was considered. For this case, Eq. (6.25) gives $\alpha = 1.6\tau_r$, where τ_r is the width of the rectangular pulse before passing through a rectangular filter of bandwidth B_s. The half-power width τ after the pulse passes through the band-limited filter is $0.625\tau_r$, so that $\alpha = 2.6\tau$.

The value of α for a perfectly rectangular pulse is finite even though the value of β for a perfectly rectangular pulse was found to be infinite. The effective time duration α will be infinite, however, for a waveform that has a perfectly rectangular frequency spectrum of width B. Such a spectrum corresponds to a $(\sin x)/x$ time waveform of infinite duration, where $x = \pi Bt$. Any practical waveform must be limited in time, and α will therefore be finite. The frequency error for a time-limited waveform with a rectangular-like spectrum may be found in a manner similar to that which was employed for finding the time-delay error of a band-limited rectangular pulse. The $(\sin x)/x$ time waveform is limited to a duration T_s just as the $(\sin x)/x$ spectral bandwidth was limited to a bandwidth B_s. The frequency error can be found from Eq. (6.11) except the roles of bandwidth and frequency are reversed. In Eq. (6.11), replace the time-delay error δT_R with δf, replace the pulse width τ_r with the bandwidth B, and replace the bandwidth B_s with the signal duration T_s.

Trapezoidal Pulse The theoretical rms error for the measurement of frequency with a trapezoidal pulse is

$$\delta f = \frac{(3\tau_p + 2t_r)^{1/2}}{2\pi\left(\dfrac{\tau_p^3}{4} + \dfrac{\tau_p^2 t_r}{2} + \dfrac{\tau_p t_r^2}{2} + \dfrac{t_r^3}{5}\right)^{1/2}\left(\dfrac{2E}{N_0}\right)^{1/2}} \qquad \text{trapezoidal pulse} \qquad \textbf{[6.26]}$$

where the rise and fall time is t_r and the width across the top of the trapezoid is τ_p. Assuming, for example, that $t_r = \tau_p/2$, the value of α would be $0.81\pi\tau_p$. When t_r is small compared to τ_p, the value of α approaches $\pi\tau_p/\sqrt{3}$, which is the value found in the above for the perfectly rectangular pulse.

Triangular Pulse This is obtained from the expression for the trapezoidal pulse by setting $\tau_p = 0$ and letting $2t_r = \tau_B$, where τ_B = width of the base of the triangular waveform. This results in

$$\delta f = \frac{(10)^{1/2}}{\pi\tau_B(2E/N_0)^{1/2}} \qquad \text{triangular pulse} \qquad [6.27]$$

Gaussian Pulse The rms error in frequency for a gaussian pulse is

$$\delta f = \frac{1.18}{\pi\tau(2E/N_0)^{1/2}} = \frac{B}{1.18(2E/N_0)^{1/2}} \qquad \text{gaussian pulse} \qquad [6.28]$$

where τ = half-power pulse width and B = half-power bandwidth.

Multiple Observations The error expressions for time delay and frequency presented here apply for a single observation. When more than one independent measurement is made, the resultant error is reduced and may be found by combining errors in the usual manner for gaussian statistics: the variance (the square of δT_R or δf) of the N independent observations is equal to $1/N$ of the variance of a single observation. Alternatively, the expression for a single pulse applies to multiple pulses if the energy E is the total energy of N pulses. The above assumes that the effective bandwidth β or the effective time-duration α remains the same for each of the N measurements.

Certainty of the Uncertainty Principle The product of the effective bandwidth and the effective time-duration α must be equal to or greater than π; that is,

$$\beta\alpha \geq \pi \qquad [6.29]$$

This relation may be derived from the definitions of β and α given by Eqs. (6.8) and (6.23) and by applying the Schwartz inequality Eq. (5.11). It is a consequence of the Fourier-transform relationship between a time waveform and its spectrum. The longer the time duration of a waveform, the narrower will be its spectrum. The wider the spectrum, the narrower will be the time waveform. Both the time waveform and its frequency spectrum cannot be made arbitrarily small simultaneously.

Equation (6.29) has sometimes been referred to as the *radar uncertainty principle* because of its supposed analogy to the important concept in quantum physics known as the Heisenberg uncertainty principle. The physics uncertainty principle states that both the position and velocity of an object (such as a subatomic particle) cannot be measured exactly at the same time.[17] Equation (6.29) actually has the opposite interpretation for radar

signals, and it is inappropriate to refer to it as a radar uncertainty principle. It follows from this equation that there is no theoretical restriction on accuracy with which a radar can simultaneously locate the position of a target and determine its velocity. The product of the rms time-delay error, Eq. (6.7), and the rms frequency error, Eq. (6.22), is

$$\delta T_R \delta f = \frac{1}{\beta\alpha(2E/N_0)} \qquad [6.30]$$

Substituting the inequality of Eq. (6.29) into the above gives

$$\delta T_R \delta f \leq \frac{1}{\pi(2E/N_0)} \qquad [6.31]$$

This states that the time delay and the frequency may be simultaneously measured to as small a theoretical error as one desires by designing the radar to yield a sufficiently large ratio of signal energy (E) to noise power per unit bandwidth (N_0), or for fixed E/N_0, to select a waveform with a large $\beta\alpha$ product. Large $\beta\alpha$ requires waveforms with long time duration and wide spectral width. In terms of range accuracy δR and radial-velocity accuracy δv_r, the expression of Eq. (6.31) can be written

$$\delta R \delta v_r \leq \frac{c\lambda}{4\pi(2E/N_0)} \qquad [6.32]$$

where λ = radar wavelength and c = velocity of propagation. This states that the shorter the wavelength, the better will be the accuracy that can be achieved in the simultaneous measurement of range and radial velocity.

There is nothing "uncertain" about the simultaneous radar measurement of range and radial velocity. The radar "uncertainly relation" and the physics uncertainty principle that describes quantum mechanical effects should not be confused. In quantum mechanics, the observer does not have control over the waveform with which the quantum particle is being observed. The radar engineer, on the other hand, can choose the value of the $\beta\alpha$ product, the signal energy E, and to some extent the noise density N_0. Any limits to classical radar measurement accuracies are practical ones.

The use of the Schwartz inequality in deriving Eq. (6.29) shows that the poorest waveform for obtaining accurate time-delay and frequency measurements simultaneously is the one for which $\beta\alpha = \pi$ (the smallest theoretical value allowed for $\beta\alpha$). This corresponds to the gaussian-shaped pulse. The quasi-rectangular pulse has $\beta\alpha = 1.22\pi$, and the trapezoidal pulse with rise time $t_r = \tau_p/2$ has $\beta\alpha = 1.4\pi$. Thus there is not much difference in the $\beta\alpha$ products of simple-shape waveforms. Large values of $\beta\alpha$ require internal modulation of the pulse to make the bandwidth much greater than the reciprocal of the pulse width. This is what is done in pulse compression waveforms, to be discussed later in this chapter (Sec. 6.5).

Angular Accuracy The expression for the theoretical accuracy with which an antenna can measure the angle of arrival follows from the above discussion of accuracy for the time-delay measurement. This is possible because of the similarity of the mathematics in

the spatial (angle) and spectral (frequency) domains. The one-dimensional electric-field strength pattern of an antenna in one plane is given in Chap. 9 as

$$g(\theta) = \int_{-D/2}^{+D/2} A(z) \exp\left(j2\pi \frac{z}{\lambda} \sin\theta\right) dz \qquad [6.33]$$

where the antenna is of length D and lies along the z-axis, $A(z)$ is the distribution of the current across the aperture (called the *aperture illumination*), λ = radar wavelength, and θ = angle measured from broadside ($\theta = 0$ is the perpendicular to the antenna). This is an inverse Fourier transform and resembles the inverse Fourier transform between the time waveform $s(t)$ and the spectrum $S(f)$.

$$s(t) = \int_{-\infty}^{\infty} S(f) \exp(j2\pi ft)\, df \qquad [6.34]$$

The antenna pattern $g(\theta)$ can be related to the time waveform $s(t)$, the aperture illumination $A(x)$ to $S(f)$, $\sin\theta$ to the time t, and the aperture coordinate z/λ can be related to the frequency f. What is learned about signals in the frequency domain often can be applied to signals in the spatial domain—and vice versa. Using the above analogies, the rms error for an angle measurement can be obtained from Eqs. (6.7) and (6.8) as

$$\delta\theta = \frac{1}{\gamma(2E/N_0)^{1/2}} \qquad [6.35]$$

where the effective aperture width γ is defined as

$$\gamma^2 = \frac{\int_{-\infty}^{\infty} (2\pi z/\lambda)^2 |A(z)|^2\, dz}{\int_{-\infty}^{\infty} |A(z)|^2\, dz} \qquad [6.36]$$

The effective aperture width is 2π times the square root of the normalized second moment of $|A(z)|^2$ about the mean value of z. The mean is at $z = 0$.

The theoretical angle-measurement error for an antenna with a uniform (rectangular) amplitude illumination across the aperture is

$$\delta\theta = \frac{\sqrt{3}\lambda}{\pi D(2E/N_0)^{1/2}} = \frac{0.628\theta_B}{(2E/N_0)^{1/2}} \qquad [6.37]$$

where the far-right expression uses the relation that the half-power beamwidth for this aperture illumination is $\theta_B = 0.88\lambda/D$. The units of $\delta\theta$ and θ_B are radians. With a cosine illumination $A(z) = \cos(\pi z/D)$ across the aperture of dimension D (where $|z| \leq D/2$), the effective aperture width γ is $1.13D/\lambda$, or $1.37/\theta_B$. A triangular illumination has $\gamma = 0.99D/\lambda$; and for a parabolic illumination, $\gamma = 0.93D/\lambda$.

Commonality of Measurements The measurements of range, angle, and radial velocity are all carried out differently in radar, but they share the concept of locating the maximum of a time waveform similar to that in Fig. 6.4. This figure might represent the radar echo time-waveform and the measurement of time delay (range); or it might represent a

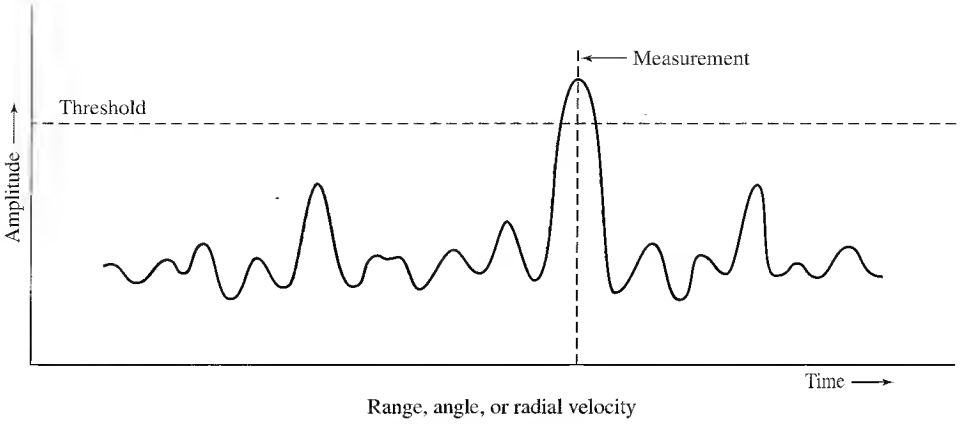


Figure 6.4 Representation of the radar measurement of range, angular location, or radial velocity by observing the waveform as a function of time, angle, or frequency.

scanning antenna pattern and the measurement of angle; or it might be the doppler frequency as observed at the output of a tunable filter that provides the radial velocity. The range, angle, or radial velocity is determined when the "signal" is a maximum.

6.4 AMBIGUITY DIAGRAM

It was mentioned in Sec. 5.2 that the output of the matched filter is the cross correlation between (1) the received signal plus noise and (2) the transmitted signal. When the signal-to-noise ratio is large (as it must be for detection), the output of the matched filter can usually be approximated by the *autocorrelation function* of the transmitted signal; that is, the noise is ignored. This assumes there is no doppler shift so that the received echo signal has the same frequency as the transmitted signal. In many radar applications, however, the target is moving so that its echo signal has a doppler frequency shift. The output of the matched filter, therefore, will not be the autocorrelation function of the transmitted signal. Instead, it must be considered as the cross correlation between the doppler-shifted received signal and the transmitted signal, with noise being ignored since the signal-to-noise ratio is assumed to be large.

The nature of the matched-filter output as a function of both time and doppler frequency is important for understanding the properties of a radar waveform, in particular its effect on measurement accuracy, target resolution, ambiguities in range and radial velocity, and the response to clutter. These aspects of the matched-filter output are examined next.

When the received echo signal is large compared to noise, the output of the matched filter [Eq. (5.17)] may be written as the following cross-correlation function:

$$\text{output of the matched filter} = \int_{-\infty}^{\infty} s_r(t)s^*(t - T_R')\, df \qquad [6.38]$$

where $s_r(t)$ is the received echo signal, $s(t)$ is the transmitted signal, $s^*(t)$ is its complex conjugate, and T'_R is the estimate of the time delay (considered a variable). Complex notation is assumed, so that the transmitted signal can be written as $u(t) \exp [j2\pi f_0 t]$, where $u(t)$ is the complex modulation function whose magnitude $|u(t)|$ is the envelope of the real signal, and f_0 is the carrier frequency. The received echo signal $s_r(t)$ is assumed to be the same as that transmitted, except for a doppler frequency shift f_d and a delay equal to the true time delay T_0. Therefore,

$$s_r(t) = u(t - T_0) \exp [j2\pi(f_0 + f_d)(t - T_0)] \quad [6.39]$$

(The change of amplitude due to the factors in the radar equation is ignored here.) With the above definitions, the output of the matched filter is

$$\text{output} = \int_{-\infty}^{\infty} u(t - T_0)u^*(t - T'_R)e^{j2\pi(f_0+f_d)(t-T_0)}e^{-j2\pi f_0(t-T'_R)}\, dt \quad [6.40]$$

For simplicity in understanding this equation, we take the origin to be the true time delay and the transmitted frequency; hence, $T_0 = 0$ and $f_0 = 0$. Then $T_0 - T'_R = -T'_R = T_R$. The output of the matched filter is then

$$\chi(T_R, f_d) = \int_{-\infty}^{\infty} u(t)u^*(t + T_R)\, e^{j2\pi f_d t}\, dt \quad [6.41]$$

A positive time delay T_R indicates a target beyond the true target time delay T_0, and a positive doppler frequency f_0 indicates an approaching target.[18,19] The squared magnitude of Eq. (6.41), $|\chi(T_R, f_d)|^2$, is called the *ambiguity function.* Its three-dimensional plot as a function of time delay T_R and doppler frequency f_d is the *ambiguity diagram.*[20]

Properties of the Ambiguity Diagram The ambiguity function $|\chi(T_R, f_d)|^2$ has the following properties:

$$\text{maximum value: } |\chi(T_R, f_d)|_{\max}{}^2 = |\chi(0,0)|^2 = (2E)^2 \quad [6.42]$$

$$\text{symmetry relation: } |\chi(-T_R, -f_d)|^2 = |\chi(T_R, f_d)|^2 \quad [6.43]$$

$$\text{behavior on } T_R \text{ axis: } |\chi(T_R, 0)|^2 = \left|\int u(t)u^*(t + T_R)\, dt\right|^2 \quad [6.44]$$

$$\text{behavior on } f_d \text{ axis: } |\chi(0, f_d)|^2 = \left|\int u^2(t)e^{j2\pi f_d t}\, dt\right|^2 \quad [6.45]$$

$$\text{volume under surface: } \int\int |\chi(T_R, f_d)|^2\, dT_R\, df_d = (2E)^2 \quad [6.46]$$

Equation (6.42) states that the maximum value of the ambiguity function occurs at the origin, which is the true location of the target when the doppler shift $f_d = 0$. Its maximum value is $(2E)^2$, where E is the energy contained in the echo signal. Equation (6.43) is a symmetry relation. Equation (6.44) is the form of the ambiguity function on the time-delay axis. It is the square of the autocorrelation function of $u(t)$. Equation (6.45) describes the behavior on the frequency axis and is the square of the inverse fourier transform of $[u(t)]^2$. The total volume under the ambiguity diagram is given by Eq. (6.46) and is a constant, also equal to $(2E)^2$. (All limits in the above equations go from $-\infty$ to $+\infty$.)

"Ideal" Ambiguity Diagram If there were no theoretical restrictions, the "ideal" ambiguity diagram would consist of a single peak of infinitesimal thickness at the origin and be zero everywhere else, as shown in Fig. 6.5. It would be an impulse function and have no ambiguities in range or doppler frequency (radial velocity). The infinitesimal (or very small) thickness at the origin would permit the time delay and/or frequency to be determined simultaneously to as high a degree of accuracy as desired. It would also permit the resolution of two very closely spaced targets and reject all clutter other than clutter at the origin. There would be no ambiguous responses. Such a highly desirable ambiguity diagram, however, is not theoretically allowed. It cannot be obtained because Eq. (6.42) requires that the maximum of the ambiguity diagram be equal to $(2E)^2$ and the volume under its surface as given by Eq. (6.46) also must be equal to $(2E)^2$.

The restrictions on the ambiguity diagram may be considered by imagining a box of sand.[21] The total amount of sand in the box is fixed, just as the volume under the ambiguity diagram is fixed at $(2E)^2$ by the signal energy E. The sand may be piled up in the center of the box (the origin of the ambiguity diagram), but its height can be no greater than $(2E)^2$. If one tries to pile sand at the center of the box in a very narrow pile (to obtain good accuracy and resolution), the sand that remains must be redistributed elsewhere in the box. Sand might then pile up in other parts of the box, which means ambiguities in range and/or doppler can result. Thus the nature of the ambiguity diagram indicates there have to be trade-offs made among the resolution, accuracy, and ambiguity.

An approximation to what might seem to be a good ambiguity diagram is shown in Fig. 6.6. The waveform does not result in ambiguities since there is but a single peak. In general, when a single peak is obtained, such as shown here, it might be so wide along the time-delay axis and the doppler-frequency axis that it might have poor accuracy and resolution. It appears that, in practice, waveforms generally have significant response somewhere in the ambiguity diagram outside the narrow region in the near vicinity of the origin. Practical waveforms do not approximate the ideal ambiguity diagram or even its more realistic version of Fig. 6.6.

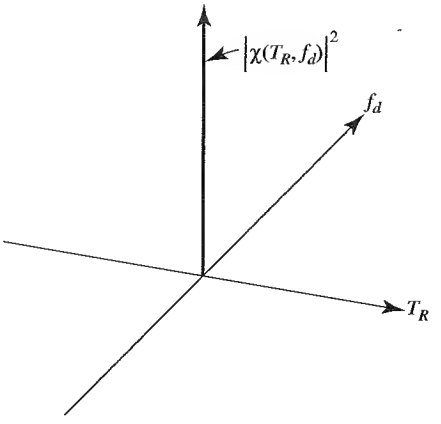


Figure 6.5 Ideal, but unattainable, ambiguity diagram.

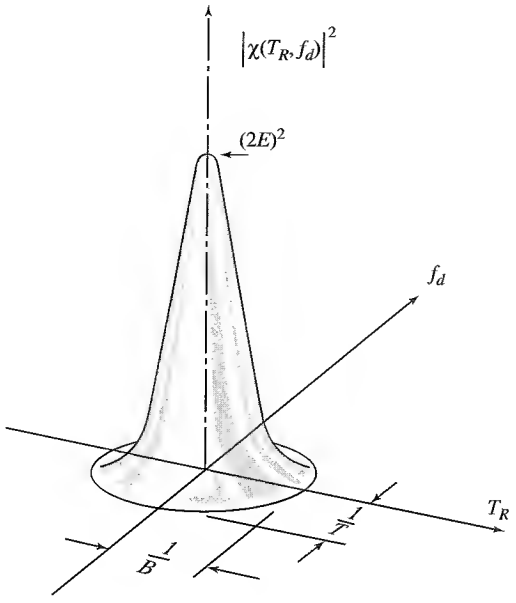


Figure 6.6 An approximation to the ideal ambiguity function with the restriction that the value at the origin is constant at $(2E)^2$ and the volume under the surface of $|\chi(T_R, f_d)|^2$ also is given by $(2E)^2$.

Single Pulse of Sinewave and the Ridge Ambiguity Diagram A computer-generated 3-D plot of $\chi|T_R, f_d|$ for a single rectangular pulse of width τ is shown in Fig. 6.7, due to A. W. Rihaczek and R. L. Mitchell. (Note that it is the magnitude of the square root of ambiguity function that is plotted here.) The triangular shape of the output of the matched filter on the time axis ($f_d = 0$) can be seen as well as the $(\sin x)/x$ shape on the frequency axis.

The essence of the information found from an ambiguity diagram can usually be obtained from simpler two-dimensional plots, as in Fig. 6.8 for (a) a long pulse of width τ and (b) for a short pulse. Shading is used to indicate the regions where $|\chi(T_R, f_d)|^2$ is

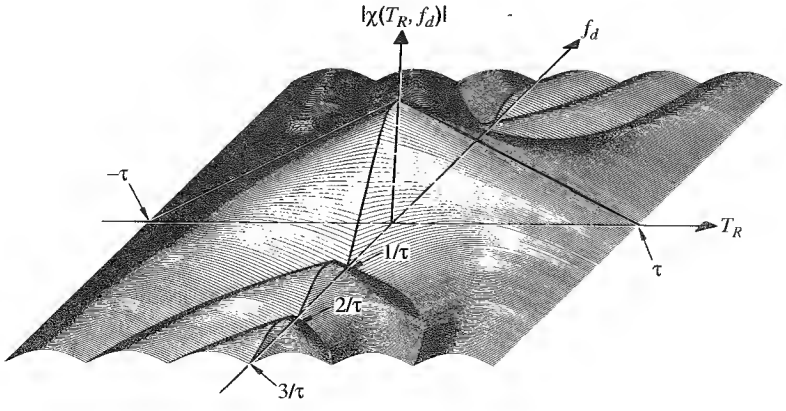


Figure 6.7 Plot of $|\chi(T_R, f_d)|$ for a simple rectangular pulse of width τ. (Due to A. W. Rihaczek and R. L. Mitchell[22])

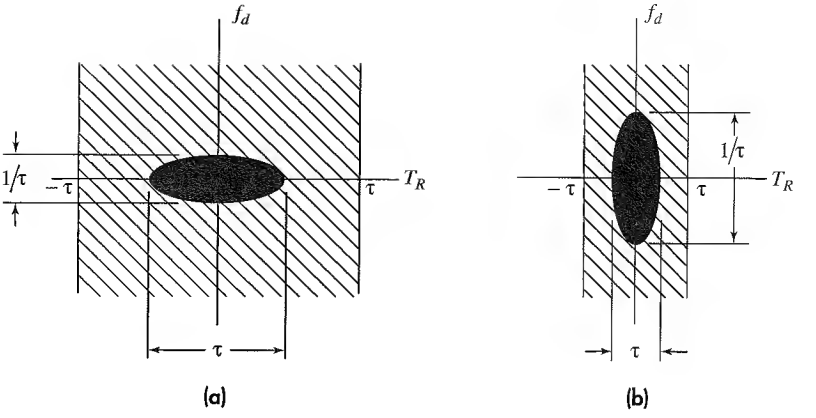


Figure 6.8 Two-dimensional representation of the ambiguity diagram for a single pulse of sinewave of width τ. (a) Long pulse; (b) short pulse.

large (completely darkened area), where it is small (lightly shaded area), and regions of zero response (no shading). The plot for a single pulse shows most of the response as a completely shaded elliptically shaped region in which $|\chi|^2$ is large. No response occurs outside the time $\pm\tau$. The time-delay measurement error is proportional to τ and the frequency measurement error is proportional to $1/\tau$. The plot of Fig. 6.8a shows the good frequency measurement accuracy and the poor time-delay accuracy of a long pulse. Figure 6.8b shows the opposite occurs for a short pulse. The ambiguity diagram indicates that as the range accuracy of a simple pulse waveform is improved, the frequency accuracy worsens and vice versa. (This follows from the box of sand analogy mentioned previously.) The short pulse is *doppler tolerant* in that a single matched filter will produce a good output if there is a significant doppler shift. That is, the output from a filter matched to a zero doppler shift will not change much when there is a doppler shift. On the other hand, a long pulse will produce greatly reduced output for a doppler-frequency shift; so it is not tolerant to changes in doppler. This is the reason why in an MTI radar a bank of doppler filters sometimes is used to cover the expected range of doppler frequencies. The ambiguity diagram of the single pulse is known as a *ridge* or *knife edge.*

Single Linear Frequency-Modulated (FM) Pulse Another example of a ridge or knife edge ambiguity diagram is that produced by linearly frequency modulating a rectangular pulse over a bandwidth B, as shown by the 2-D plot of Fig. 6.9. The pulse width T is large compared to $1/B$. The frequency modulation increases the spectral bandwidth of the pulse so that $BT >> 1$. The ridge is at an angle determined by the slope B/T. The time-delay measurement accuracy is proportional to $1/B$ and the frequency (if it can be measured) has an accuracy proportional to $1/T$. Since the pulse width T and the bandwidth B can be chosen independent of one another, the time-delay and frequency accuracies are independent of the other. This is unlike the time-delay and frequency accuracies of the simple, unmodulated rectangular pulse.

Periodic Pulse Train Consider, as shown in Fig. 6.10a, a series of five pulses, each of width τ, pulse repetition period T_p, and total duration T_d. The 2-D representation of its

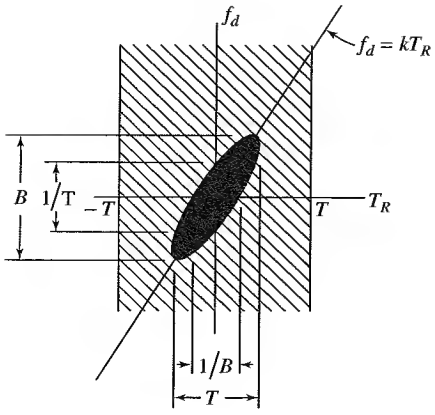
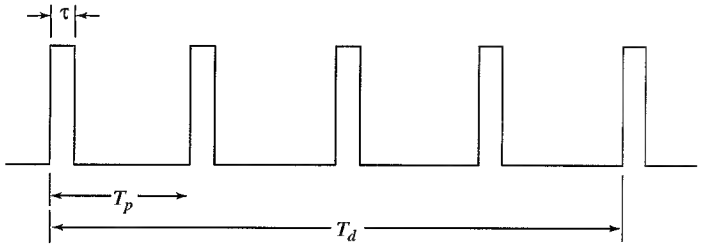
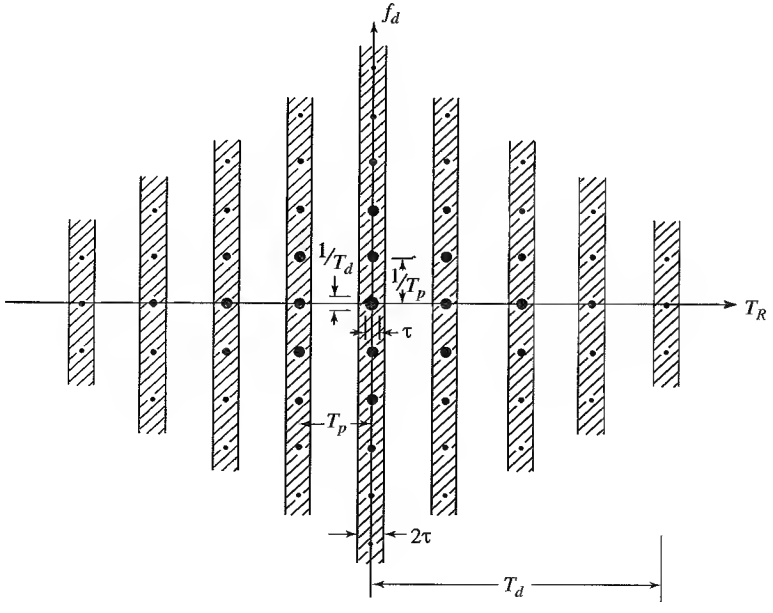


Figure 6.9 Two-dimensional representation of the ambiguity diagram for a single linear frequency-modulated pulse of width T and bandwidth B.



(a)



(b)

Figure 6.10 (a) Video pulse train of five pulses; (b) two-dimensional representation of the ombiguity diagrom for (o).

ambiguity diagram is shown in Fig. 6.10b. This type of ambiguity function is called a *bed of spikes*. Throughout this ambiguity diagram there are many possible ambiguities in range and doppler (blind speeds) such as were encountered before in this text. Ambiguities are a consequence of discontinuous waveforms, which is why they do not appear with a single pulse. It might seem at first glance that the many ambiguities produced by a pulse train result in it being a poor radar waveform. As seen in Chap. 3, a pulse train is widely used in pulse doppler and MTI radars, and it is often possible to adequately unravel, avoid, or (sometimes) ignore the ambiguities.

Examine the center (the origin) of the ambiguity diagram shown in Fig. 6.10b. The spike at the origin is of dimension τ on the time axis and $1/T_d$ on the frequency axis. The time-delay measurement accuracy (determined by the pulse width) and the frequency accuracy (determined by the duration of the signal) can be selected independently. If the pulse repetition period T_p is such that no radar echoes are expected with a time delay greater than T_p and no doppler-frequency shifts are expected greater than $1/T_p$, then the effective ambiguity diagram reduces to just a single spike at the origin whose dimensions are determined by τ and T_d. The pulse train, therefore, can be a good radar waveform. Ambiguities that occur can be resolved with different pulse repetition frequencies (Secs. 2.10 and 3.9). The fact that many radars employ this type of waveform attests to its usefulness far better than any analysis based on its ambiguity diagram.

Although this discussion applies to a pulse train, the signal processing employed with such waveforms is almost always different from that assumed for Fig. 6.10. The ambiguity function of a pulse train is the output of a matched filter designed for the *entire* pulse train. This is why the ambiguity diagram of Fig. 6.10 for five pulses has an output (in time) consisting of nine pulses of different amplitudes when the input is five equal amplitude pulses (the autocorrelation function of five pulses). Pulse radars, however, are not usually designed with a matched filter for N pulses. Instead, a pulse radar usually uses a filter matched to a single pulse and then integrates the N pulses.

Noiselike Waveforms and the Thumbtack Ambiguity Diagram A noise waveform has an ambiguity function, called a thumbtack, similar to that of Fig. 6.11. Pure noise waveforms are seldom employed in radar; but constant-amplitude noiselike waveforms have been used to produce a thumbtack ambiguity diagram. Examples include nonmonotonic frequency modulation and pseudorandom variations of phase or frequency. These are discussed later in this chapter under the topic of pulse compression.

The thumbtack has the advantage that the time-delay and frequency measurement accuracies are independently determined, respectively, by the bandwidth of the modulation and the duration of the pulse. There are no apparent ambiguities and it resembles the approximation of the ideal ambiguity function of Fig. 6.6, except for the plateau on which the main response rests. This plateau extends over a dimension $2T$ along the time axis and $2B$ along the frequency axis. When the product BT is large, the volume under the peak response is small compared to the total volume, and almost all of the volume is in the plateau. Consequently, the average height of the plateau is approximately E^2/BT for large BT (based on Eq. 6.46).

The thumbtack ambiguity diagram, as illustrated in the sketch of Fig. 6.11, appears attractive; but it is an overly simplified sketch that can be misleading when practical

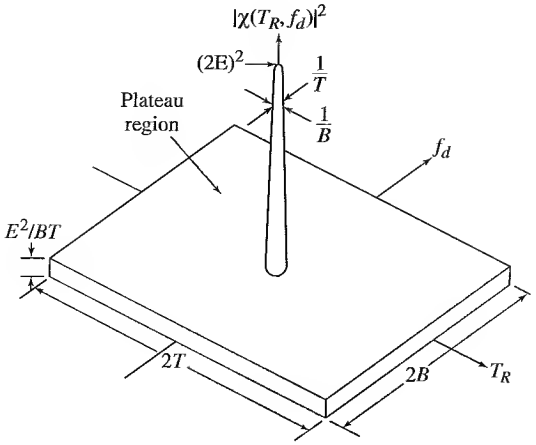


Figure 6.11 Ideal representation of the thumbtack ambiguity diagram as might be produced by a naiselike wavefarm or a pseudarandom caded pulse wavefarm (details af the sidelobe structure are not shawn.)

waveforms are examined. The plateau is not uniform as shown. Most of the volume under the ambiguity function is confined to the intervals $|T_R| < T/2$ and $|f_d| < B/2$.[23] Furthermore, when there is a doppler-frequency shift, the sidelobes can be relatively high and can cause false responses. There are some waveforms, as will be mentioned later, that can produce low sidelobes when $f_d = 0$; but such waveforms should be looked at with caution when good doppler filtering is required without false responses, as for the detection of moving targets.

Waveform Design and the Ambiguity Diagram The waveform transmitted by a radar can affect (1) target detection, (2) measurement accuracy, (3) resolution, (4) ambiguities, and (5) clutter rejection. The ambiguity diagram may be used to assess qualitatively how well a particular waveform achieves these capabilities. Each of the above five capabilities will be briefly discussed.

Detection It was said in Sec. 5.2 that if the receiver is designed as a matched filter, the output peak-signal-to-mean-noise ratio (which is related to the ability to detect a target) depends only on the ratio of the received signal energy E and the receiver noise power per unit bandwidth N_0. The requirements for detection do not place any demands on the *shape* of the transmitted waveform except that (1) the waveform be practical to generate and radiate, and (2) the matched filter required for the waveform be practical to achieve. The ambiguity diagram, therefore, is seldom used to assess the detection capability of a particular waveform, except to note if the signal contains sufficient energy.

Accuracy The accuracy with which the range and radial velocity can be measured is indicated by the main response at the origin. The width along the time axis determines the range (time delay) accuracy, and the width along the frequency axis determines the radial-velocity accuracy.

Resolution The width of the central response also determines the resolution ability of a radar waveform in range and radial velocity. In order to resolve closely spaced targets, the central response must be isolated. There cannot be extraneous peaks near the main response that would mask the echo from a nearby target.

Ambiguity Ambiguities occur in radar measurements when the waveform is not continuous. For example, a pulse train is not continuous; hence, such a waveform can produce ambiguities in range (time delay) and in velocity (frequency). An ambiguous measurement is one in which there is more than one choice available, but only one is correct. Ambiguities appear in the ambiguity diagram as additional high responses similar in magnitude to the peak response at the origin. The correct response from the target cannot be readily recognized from these additional responses without taking some action to resolve the correct from the incorrect values.

The name *ambiguity function* for $|\chi(T_R, f_d)|^2$ can be misleading since this function describes more about the character of a waveform than just the ambiguities it produces. Woodward[24] coined the name for an entirely different reason than the ambiguities associated with discontinuous waveforms. The reader is advised not to be distracted by trying to understand the ambiguous use of the term "ambiguity" as the name for the function $|\chi(T_R, f_d)|^2$.

Clutter Attenuation Resolution in range and velocity can enhance the target signal echo relative to nearby distributed clutter echoes. The ambiguity diagram can indicate the ability of a waveform to reject clutter by superimposing the locations of the clutter echoes (as stored in a *clutter map*) on the T_R,f_d plane of the ambiguity diagram. If the radar is to have good clutter rejection, the ambiguity diagram should have little or no response in regions of high clutter echoes. The short-pulse waveform, for example, will reduce stationary clutter that is extensive in range since the ambiguity function for this waveform has little response on the time (range) axis.

Waveform Synthesis It is reasonable to ask if radar waveforms can be obtained by first defining a desired ambiguity diagram and then synthesizing the signal that yields this ambiguity diagram. Synthesis has not proven successful in the past. Instead, it is more usual to compute the ambiguity diagrams of various waveforms and determine which of them have suitable properties for the intended application.

6.5 PULSE COMPRESSION

High range-resolution, as might be obtained with a short pulse, is important for many radar applications, as indicated by the list of capabilities in Table 6.2. There can be limitations, however, to the use of a short pulse. Since the spectral bandwidth of a pulse is inversely proportional to its width, the bandwidth of a short pulse is large. Large bandwidth can increase system complexity, make greater demands on the signal processing, and increase the likelihood of interference to and from other users of the electromagnetic

Table 6.2 Capabilities of short-pulse, high range-resolution radar

Range resolution. Usually easier to separate (resolve) multiple targets in range than in angle.
Range accuracy. A radar capable of good range resolution is also capable of good range accuracy.
Clutter reduction. Increased target-to-clutter ratio is obtained by reducing the amount of distributed clutter with which the target echo signal must compete.
Interclutter visibility. With some types of "patchy" land and sea clutter, a high-resolution radar can detect moving targets in the clear areas between the clutter patches.
Glint reduction. Angle and range tracking errors introduced by a complex target with multiple scatterers are reduced when high range-resolution is employed to isolate (resolve) the individual scatterers that make up the target.
Multipath resolution. Range resolution permits the separation of the desired target echo from the echoes that arrive at the radar via scattering from longer propagation paths, or multipath.
Multipath height-finding. When multipath due to scattering of radar energy from the earth's surface can be separated from the direct-path signal by high range-resolution, target height can be determined without a direct measurement of elevation angle.
Target classification. The range, or radial, profile of a target in some cases can provide a measure of target size in the radial dimension. From the range profile one might be able to sort one type of target from another based on size or distinctive profile, especially if the cross-range profile is also available.
Doppler tolerance. With a short-pulse waveform, the doppler-frequency shift from a moving target will be small compared to the receiver bandwidth; hence, only a single matched filter is needed for detection, rather than a bank of matched filters each tuned for a different doppler shift.
ECCM. A short-pulse radar can negate the effects of certain electronic countermeasures such as range-gate stealers, repeater jammers, and decoys. The wide bandwidth of the short-pulse radar can, in principle, provide some reduction in the effects of broadband noise jamming and reduce the effectiveness of some electronic warfare receivers and their associated signal processing.
Minimum range. A short pulse allows the radar to operate with a short minimum range. It also allows reduction of blind zones (eclipsing) in high-prf radars.

spectrum. Another limitation is that in some high-resolution radars the limited number of resolution cells available with conventional displays might result in overlap of nearby echoes when displayed, which results in a collapsing loss (Sec. 2.12) if the detection decision is made by an operator. Wide bandwidth can also mean less dynamic range in the receiver because receiver noise power is proportional to bandwidth. Also, a short-pulse waveform provides less accurate radial velocity measurement than if obtained from the doppler-frequency shift. It spite of such limitations, the short-pulse waveform is used because of the important capabilities it provides.

A serious limitation to achieving long ranges with short-duration pulses is that a high peak power is required for a large pulse energy. The transmission line of a high peak power radar can be subject to voltage breakdown (arc discharge), especially at the higher frequencies where waveguide dimensions are small. If the peak power is limited by breakdown, the pulse might not have sufficient energy. Consider, for example, a more or less conventional radar with a pulse width of one microsecond and one megawatt peak power, as might be found in a medium-range air-surveillance radar. In this example, the energy contained in a single pulse is one joule. (The energy per pulse and the number of pulses

integrated determine the detectability of a target.) A one microsecond pulse has a range resolution of 150 m. If it were desired to have a resolution of 15 cm (one-half foot), the pulse width would have to be reduced to one nanosecond and the peak power increased to *one gigawatt* (10^9 W) in order to maintain the same pulse energy of one joule. This is an unusually large peak power that cannot be propagated without breakdown in the usual types of transmission lines employed at microwave radar frequencies.

A short pulse has a wide spectral bandwidth. A long pulse can have the same spectral bandwidth as a short pulse *if* the long pulse is modulated in frequency or phase. (Amplitude modulation can also increase the bandwidth of a long pulse, but is seldom used in radar because it can result in lower transmitter efficiency.) The modulated long pulse with its increased bandwidth B is compressed by the matched filter of the receiver to a width equal to $1/B$. This process is called *pulse compression*. It can be described as the use of a long pulse of width T to obtain the resolution of a short pulse by modulating the long pulse to achieve a bandwidth $B >> 1/T$, and processing the modulated long pulse in a matched filter to obtain a pulse width $\tau \approx 1/B$. Pulse compression allows a radar to simultaneously achieve the energy of a long pulse and the resolution of a short pulse without the high peak power required of a high-energy short-duration pulse. It is used in high-power radar applications that are limited by voltage breakdown if a short pulse were to be used. Airborne radars might experience breakdown with lower voltages than ground-based radars, and might be candidates for pulse compression. It is almost always used in high-power radars with solid-state transmitters since solid-state devices, unlike vacuum tubes, have to operate with high duty cycles, low peak power, and pulse widths much longer than normal. Pulse compression is also found in SAR and ISAR imaging systems to obtain range resolution comparable to the cross-range resolution.

The pulse compression ratio is defined as the ratio of the long pulse width T to the compressed pulse width τ, or T/τ. The bandwidth B and the compressed pulse width τ are related as $B \approx 1/\tau$. This would make the pulse compression ratio approximately BT. If amplitude weighting of the received waveform (but not the transmitted waveform) is used to reduce the time sidelobes of the linear-FM waveform (as will be discussed later in this section), the pulse compression ratio defined as BT usually is a little larger than T/τ. It is better, therefore, to define the pulse compression ratio by T/τ (the ratio of the before and after pulse widths) rather than the bandwidth-time product BT when weighting on receive is used. (In spite of this caution, the pulse compression ratio is often given as BT in this text as well as in other radar literature. The reader should be aware of which value is used, especially with linear-FM waveforms.) The pulse compression ratio in practical radar systems might be as small as 10 (although 13 is a more typical lower value) or greater than 10^5. Values from 100 to 300 might be considered typical.

There are two ways to describe the operation of a pulse compression radar. One is based on the ambiguity function of Sec. 6.4. A long pulse is modulated to increase its bandwidth. On reception the modulated long-pulse echo signal is passed through the matched filter. Its resolution in range can be found from examination of the ambiguity diagram. The constant-amplitude linear-FM pulse, whose ambiguity diagram was shown in Fig. 6.9, is an example of a widely used pulse compression waveform. Its ambiguity diagram shows that the long pulse of width T provides a compressed pulse width equal to $1/B$.

The other method for describing pulse compression is based on how linear-FM pulse compression was presented in the original patent of R. H. Dicke,[25] before the concept of the ambiguity function was known. The modulation applied to a long pulse can be considered as providing distinctive "marks" in either the frequency or the phase along the various portions of the pulse. For instance, the changing frequency of a linearly frequency-modulated pulse is distributed along the pulse so that each small segment of the pulse corresponds to a different frequency. By passing the modulated pulse through a dispersive delay line whose delay time is a function of frequency, each part of the pulse experiences a different delay, so that it is possible for the trailing edge of the pulse to be speeded up in a dispersive* delay line and the leading edge slowed down so that they "come together" to effect a compression of the long pulse.

There have been two general classes of pulse-compression waveforms used in the past. The most popular has been the *linear FM* (also known as chirp) to be discussed next. *Binary phase-coded* pulses is the other type. Also to be mentioned here are *polyphase codes* whose phase quantizations are less than π radians; *Costas codes* in which the frequencies of the subpulses are changed in a prescribed fashion; *nonlinear FM*, *nonlinear binary phase-coding*; *complimentary codes* that in principle produce zero time-sidelobes; and codes with very low or no sidelobes that require amplitude modulation of the subpulses on transmit. No one type of pulse-compression waveform can do everything that might be required; but linear FM probably has been the most widely used.

Linear Frequency Modulation (LFM) Pulse Compression The basic concept of the linear frequency-modulated pulsed compression radar was described by R. H. Dicke in a patent filed in 1945 and issued in 1953.[25] Figure 6.12, which is derived from Dicke's patent, is a block diagram of such a radar. It is similar to the block diagram of a conventional radar except that the transmitter is shown here as being frequency modulated and the receiver

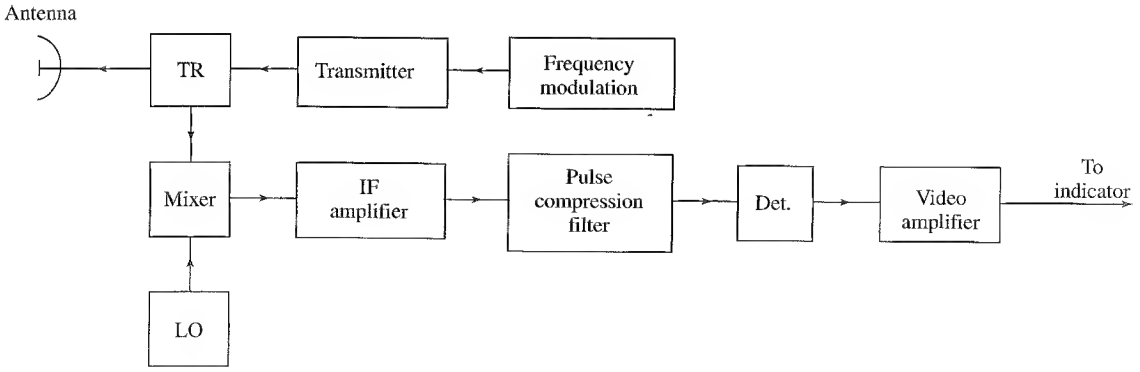


Figure 6.12 Block diagram of an FM pulse compression radar.

* "Dispersive" in this case means that the velocity of propogotion is o function of frequency.

contains a pulse-compression filter (which is identical to a matched filter). More usually, however, the FM waveform is generated at low power and amplified by a power amplifier rather than by frequency modulating a power oscillator as indicated in this figure. The transmitted waveform (Fig. 6.13a) consists of a rectangular pulse of constant amplitude A and duration T. The frequency increases linearly from f_1 to f_2 over the duration of the pulse (Fig. 6.13b). This is sometimes known as an *up-chirp*. Alternatively, the frequency could just as well decrease with time, and it is then called *down-chirp*. The time waveform is represented in Fig. 6.13c. On reception, the frequency-modulated signal is passed through the pulse-compression filter, which is a delay line whose velocity of propagation is proportional to frequency. It speeds up the higher frequencies at the trailing edge of the pulse relative to the lower frequencies at the leading edge so as to compress the signal to a width $1/B$, where $B = f_2 - f_1$ (Fig. 6.13d). The pulse compression filter is a matched filter; hence, its output envelope (neglecting noise) is the autocorrelation function of the input. In this case, the output is proportional to $(\sin \pi Bt)/\pi Bt$.[26,27] The peak power of the pulse is increased by the pulse compression ratio $BT \approx T/\tau$ after passage through the filter.

The ambiguity diagram for a linear-FM pulse-compression waveform (Fig. 6.9) shows that a large doppler shift in the echo signal can result in the indicated range not being the true range. This is known as *range-doppler coupling*. In many cases, the range error due to the doppler shift is small and can be tolerated. If the range error is large, the effect of the doppler can be removed by averaging the two range indications obtained with both a rising FM (up-chirp) and a falling FM (down-chirp) similar to the manner in which the range error due to doppler is eliminated in an FM-CW radar with triangular modulation.[28]

Reduction of Time (Range) Sidelobes The $(\sin \pi Bt)/\pi Bt$ envelope out of a matched filter when the input is a linear-FM sinewave has relatively high peak time-sidelobes of -13.2 dB adjacent to the main response. This is usually not acceptable since high sidelobes can be mistaken for targets or can mask nearby weaker targets. The time sidelobes can be reduced by transmitting a pulse with nonuniform amplitude; that is, by amplitude weighting the long pulse over its duration T. (This is similar to *tapering* the aperture illumination of an antenna, Sec. 9.3, or *windowing* to reduce the spectral sidelobes of a digital filter.) Unfortunately, it is often not practical in high-power radar to have a transmitted waveform whose amplitude varies over the pulse duration. High-power transmitters such as klystrons, traveling wave tubes, and crossed-field devices should be operated saturated to obtain maximum efficiency. They don't like to be operated with amplitude modulation and they should be either full-on or full-off. Solid-state transmitters can have a linear input-output relation and be amplitude modulated if operated Class-A; but they are almost always operated Class-C because of the much higher efficiency of Class-C.[29] It is seldom that high-power microwave radar transmitters are operated with a deliberate change in amplitude over the pulse duration.

As a compromise, the time sidelobes with a linear-FM pulse-compression signal usually are reduced by applying amplitude weighting on receive (to the dispersive delay-line pulse compression filter) and not on transmit. Since the pulse compression filter is the matched filter, applying the weighting only on receive results in a mismatched filter and a loss in signal-to-noise ratio.[30] This is the price paid to reduce the time sidelobes. Table 6.3 gives examples of weightings, the peak sidelobe that results, and other properties of

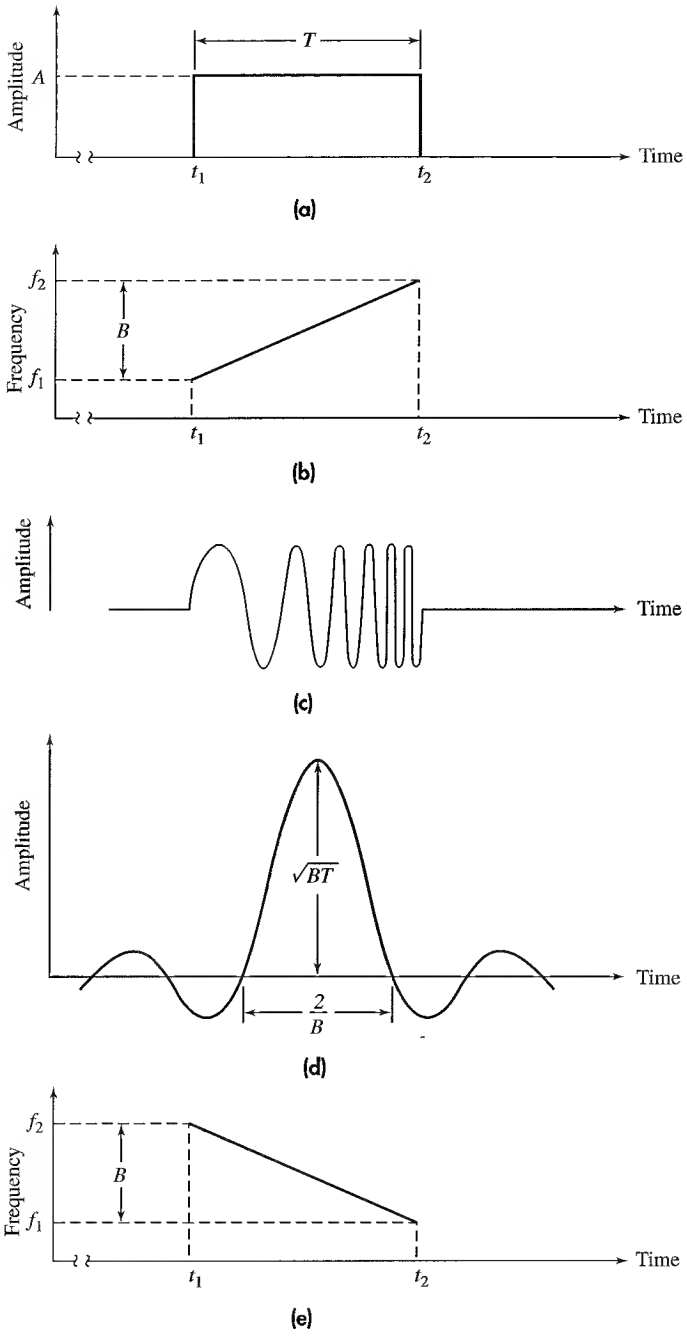


Figure 6.13 Linear FM pulse compression. (a) Transmitted waveform; (b) frequency of the transmitted waveform as a function of time; (c) representotion of the linear FM waveform; (d) theoretical output from the pulse-compression filter.

Table 6.3 Properties of weighting functions to reduce time sidelobes

Weighting Function	Peak sidelobe dB	Loss dB	Mainbeam width (relative)
Uniform	−13.2	0	1.0
$0.33 + 0.66 \cos^2(\pi f/B)$	−25.7	0.55	1.23
$\cos^2(\pi f/B)$	−31.7	1.76	1.65
$0.16 + 0.84 \cos^2(\pi f/B)$	−34.0	1.0	1.4
Taylor ($\bar{n} = 6$)	−40.0	1.2	1.4
$0.08 + 0.92 \cos^2(\pi f/B)$ (Hamming)	−42.8	1.34	1.5

the output waveform. The mismatched-filter loss can generally be kept to about 1 dB when the peak sidelobe level is reduced to 30 dB below the peak response. It is a loss that is tolerated in order to achieve lower time-sidelobe levels.

It is possible to have no theoretical loss in signal-to-noise ratio and still achieve low time-sidelobes with a uniform-amplitude transmitted waveform if a nonlinear frequency-modulated waveform is used, as discussed later in this section.

Stretch[31,32] This is a variant of linear-FM pulse compression in which narrow bandwidth processing can be employed if the extent of the range interval where high-range resolution occurs can be reduced. Stretch starts with a linear-FM waveform that is of much smaller bandwidth B_1 than required for the desired resolution. Before being transmitted, it is heterodyned (mixed) with a wideband linear FM of bandwidth B_2 to produce a signal with the desired bandwidth $B_1 + B_2$, which is radiated. On receive, the signal of bandwidth $B_1 + B_2$ is heterodyned again with the linear-FM sweep of wide bandwidth B_2. The difference signal has the narrow bandwidth B_1 and is processed as a normal (narrowband) pulse-compression signal. The heterodyne operation on receive results in a time expansion by a factor $\alpha = (B_1 + B_2)/B_1$ (time expansion means smaller bandwidth). The range resolution of a signal, however, corresponds to a signal with bandwidth $B_1 + B_2 = B$. The processing of the linear-FM signal and its generation are both done with narrow bandwidth circuitry of bandwidth $B_1 = B/\alpha$. The reduced bandwidth in waveform generation and processing is the advantage of Stretch.

Because of the "stretching" of the waveform during processing, the range interval over which high resolution is obtained is $1/\alpha$ that which would normally be obtained by a conventional linear-FM radar of bandwidth B. One trades processing bandwidth for the size of the high-resolution range interval. This capability might not be appropriate for a surveillance radar that has to look everywhere in range, but it is well suited to a tracking radar or a high range-resolution radar used for target classification.

The Stretch waveform was used in the Cobra Dane (AN/FPS-108), a large L-band phased array radar located in the Aleutian Islands of Alaska.[33] Its purpose was to gather information about Soviet/Russian intercontinental ballistic missiles (ICBM). It also has been used for space surveillance and ICBM warning. It performs high range-resolution

observation of ICBM reentry vehicles at ranges of about 1000 nmi. For this purpose it uses a 1000 μs duration linear-FM wideband waveform with 200 MHz bandwidth extending from 1175 to 1375 MHz. Its time-bandwidth product is 200,000, a relatively large value. Stretch processing examines a preselected range interval 250 ft in extent. The inherent range resolution is 2.5 ft, but the actual resolution is degraded to 3.75 ft since Taylor amplitude-weighting is employed on receive to reduce its time sidelobes to -30 dB.

Linear FM Pulse-Compression Filters In the first edition of the *Radar Handbook,* which appeared in 1970, the chapter on pulse compression[34] listed nine different devices that might be used as the pulse compression filter for linear FM. Since that time, the practice has narrowed to mainly two choices: the surface acoustic wave (SAW) dispersive delay line and digital processing. The 1970 edition briefly discussed the SAW device (then called a *surface-wave dispersive delay line*) but it didn't mention digital processing for pulse compression. Digital processing is usually preferred when the A/D converter can provide the very wide bandwidths required of high-resolution pulse-compression radar. The SAW device has been the method of choice, however, when large-bandwidth signal processing cannot be obtained digitally. Digital methods are therefore used when they are applicable and the analog SAW delay line is used for very high resolution when digital methods cannot compete.

There are two general technical areas of application for pulse compression. One is when the compressed pulse is very small, perhaps of the order of a nanosecond, but the original uncompressed pulse is of conventional width, perhaps a microsecond. High resolution might be of interest for synthetic-aperture imaging radars (SAR), radars designed to detect a person swimming in the water, or for radars required to recognize one class of ship from another based on inverse synthetic aperture radar. SAW devices have been appropriate as the pulse compression filter for such applications. The other area is when one starts with a long pulse, say a few hundred microseconds, and compresses it to a conventional width, for example, of the order of a microsecond. Digital processing can be used in such situations. An important example is when a solid-state transmitter is used that requires long pulses for efficient operation.

SAW Devices[35–37] The concept of a surface acoustic wave delay line is shown schematically in Fig. 6.14. It consists of a very smooth piezoelectric substrate, such as a thin slice of quartz, lithium niobate, or lithium tantalate. The function of the piezoelectric substrate is to support propagation of acoustical waves along the surface. The low velocity of acoustic waves (approximately 3500 m/s) compared to electromagnetic waves means that significant delay times can be achieved with a relatively small device. The input/output devices arranged on the surface are known as *interdigital transducers* (IDT). They are metallic thin films, usually aluminum, that are deposited on the substrate using photolithographic methods. The transducers are the means by which electrical signals are converted to acoustical signals, and vice versa. They determine the impulse response of the SAW delay line and the length of the IDT determines the duration of the signal. Since the IDT can launch waves in both the forward and back directions, the acoustic energy propagating in the opposite direction has to be attenuated so that it is not reflected and cause interference. One method for attenuating the unwanted signal is to use acoustic absorbers located at each end of the device, as indicated in Fig. 6.14.

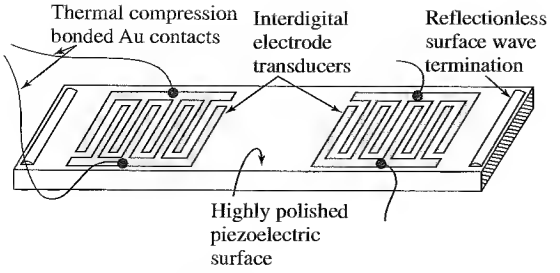


Figure 6.14 Schematic of a simple surface acoustic wave (SAW) delay line.
(From T. W. Bristol: Surface Acoustic Wave Devices—Technology and Applications, IEEE 1976 WESCON, paper 245/1. Courtesy IEEE.)

Efficient electric-to-acoustic coupling occurs when the comb fingers, or electrodes, of the IDT are spaced one-half wavelength of the acoustic signal that propagates along the substrate material. The frequency response of the delay line depends, therefore, on the spacings between the electrodes. A dispersive delay line whose delay is a function of frequency can be obtained with varying electrode spacing as illustrated in Fig. 6.15a or b. The SAW delay-line configuration in 6.15a is called an *in-line single dispersive chirp filter.* The received linear-FM signal sketched below the delay line is applied to the broadband IDT on the left. The output from the dispersive IDT structure on the right is the compressed pulse. (Either IDT can be used as the input or output since the SAW delay line is reciprocal.)

The *in-line double dispersive chirp filter* of Fig. 6.15b is capable of larger *BT* products (50 to 1000) than the single configuration of 6.15a.[35] The *slanted-array compressor* (SAC) sketched in Fig. 6.15c also is capable of large pulse compression ratios. An advantage the inclined IDTs have over the in-line configurations is that they avoid the distortion that occurs when the low-frequency components have to propagate under the high-frequency electrodes, and vice versa. Another advantage of the SAC configuration is that corrections in the phase characteristic of the device can be made by inserting a "phase plate" between the input and output IDTs. An example of a SAC linear-FM filter described by Cambell[37] (but attributed to S. Jen and C. F. Shaffer) operated at a center frequency of 1.4 GHz with a bandwidth of 1.1 GHz. This is a relative bandwidth of almost 80 percent, and can be called ultrawideband. The uncompressed pulse width was 0.44 μs, the compressed pulse was about 0.9 ns, and the bandwidth-time product was 484. Amplitude weighting was used to reduce the time sidelobes to 26.8 dB below the peak.

The *reflective-array compressor* (RAC) shown schematically in Fig. 6.15d is another form of SAW device. It is capable of larger pulse compression ratios (500 to 10,000) than other configurations. Shallow grooves are etched in the delay path that result in reflections to form a delay that depends on frequency. The structure is less sensitive to fabrication tolerances than conventional transducers. Compensation for phase errors can also be inserted between the oblique-angle grooves, as can attenuation for weighting the amplitude.

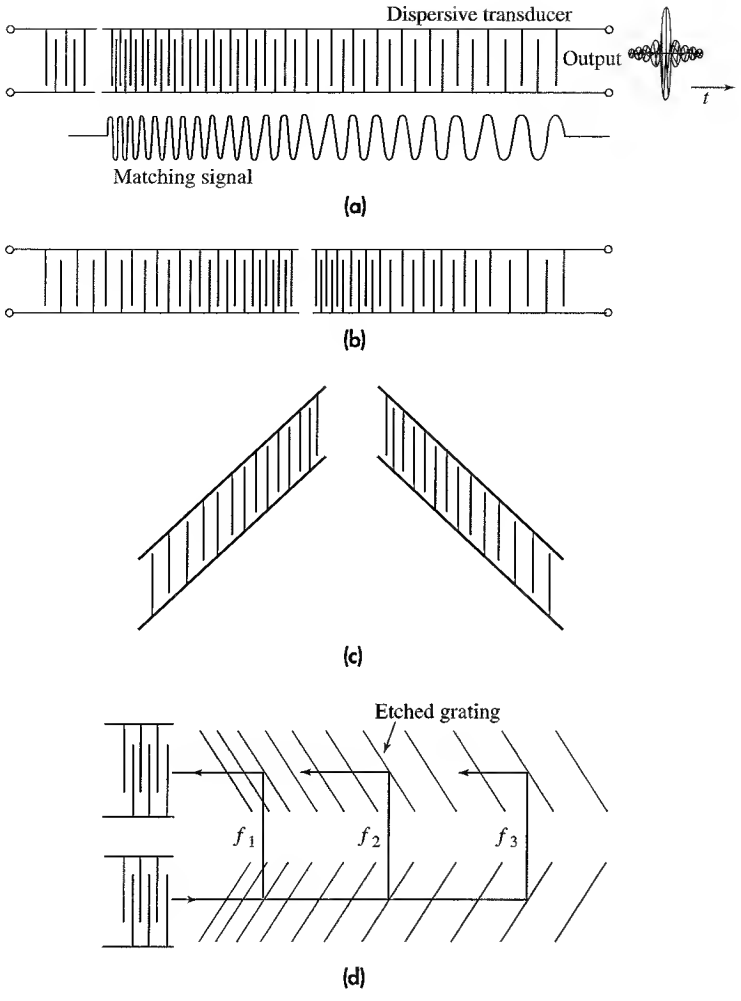


Figure 6.15 Configurations of interdigital transducers for linear-FM pulse-compression dispersive delay lines. (a) In-line single dispersive SAW delay line; (b) in-line double dispersive SAW delay line with dispersion in both transducers; (c) slanted-array compressor (SAC); and (d) reflective-array compressor (RAC). (Figs. (a), (b), and (d) are from Maines and Poige,[36] courtesy of Proc. IEEE.)

In summary, linear-FM SAW devices might have pulse compression ratios (BT) greater than 10,000, uncompressed pulse durations up to 150 μs, bandwidths greater than 1 GHz, insertion loss from 20 to 60 dB, and operate at center frequencies from 100 MHz to 1.5 GHz. They are usually found in the IF portion of the radar receiver. SAW devices provide wide bandwidth pulse-compression filters in small size packages. They are readily reproducible, easy to manufacture in large quantities, and of relatively low unit cost.

A device related to SAW devices is the IMCON, which is a reflection-mode delay line configured similar to the RAC SAW device, but fabricated on steel acoustic media. IMCONs are more suited for narrow bandwidths and low center frequencies (below 30 MHz) and when uncompressed pulse durations greater than 50 μs are desired. Bandwidths are from 0.5 to 12 MHz. Pulse durations of 600 μs are possible and units can be cascaded to obtain longer pulse durations.[38] For example, a pulse duration of 10 ms was obtained

by cascading 18 IMCONs with amplifiers. It operated at a center frequency of 7.5 MHz and bandwidth 2.5 MHz. The pulse compression ratio was 25,000 and its dimensions were 15 by 15 by 12 inches.[39]

Amplitude Weighting Amplitude shaping of the frequency response of a SAW filter can be obtained by the amount of overlap of the electrodes of the IDT, as in Fig. 6.16. This is sometimes called *apodization.*

Digital Processing for Frequency-Modulated Pulse Compression The linear-FM pulse-compression waveform, as well as most other pulse compression waveforms, can be processed and generated at low power levels by digital methods when A/D converters are available with the required bandwidth and number of bits.[40,41] Digital methods are very stable and can handle long-duration waveforms. The same basic digital system implementation can be used when one wishes to employ multiple bandwidths and pulse durations, different types of pulse compression modulations, good phase repeatability, low time-sidelobes, or when flexibility is desired in waveform selection.

Generation of the Pulse-Compression Waveform The analog and digital methods used to obtain the linear-FM pulse-compression filter can also be applied to generate the transmitted waveform. The waveforms may be generated passively or actively. An example of the former is the SAW device; an example of the latter is a voltage-controlled oscillator. Other examples of each can be found in Ref. 40.

Single Transmit-Receive Filter It will be recalled (Sec. 5.2) that the impulse response of a matched filter is the time inverse of the signal for which it is matched. If an up-chirp linear-FM is transmitted, the impulse response of its matched filter will be a down-chirp. To employ the same filter for both transmit and receive, the local oscillator frequency of the radar superheterodyne receiver should be greater than the frequency of the received signal. When the difference signal is taken from the mixing operation, the linear-FM waveform will be inverted; that is, an up-chirp becomes a down-chirp, which is the signal required for the matched filter when a transmitted up-chirp is generated using the same matched filter.

Binary Phase-Coded Pulse Compression Changes in phase can also be used to increase the signal bandwidth of a long pulse for purposes of pulse compression. A long pulse of

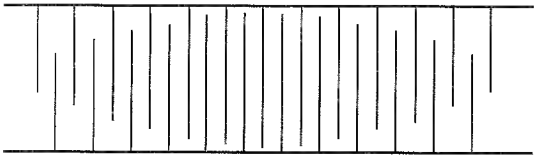


Figure 6.16 Interdigital transducer (nondispersive) showing overlap of the comb fingers to provide an amplitude weighting along the pulse.

duration T is divided into N subpulses each of width τ. An increase in bandwidth is achieved by changing the phase of each subpulse (since the rate of change of phase with time is a frequency). A common form of phase change is *binary phase coding*, in which the phase of each subpulse is selected to be either 0 or π radians according to some specified criterion. If the selections of the $0, \pi$ phases are made at random, the waveform approximates a noise-modulated signal and has a thumbtack-like ambiguity function as in Fig. 6.11. The output of the matched filter will be a compressed pulse of width τ and will have a peak N times greater than that of the long pulse. The pulse compression ratio equals the number of subpulses $N = T/\tau \approx BT$, where the bandwidth $B \approx 1/\tau$. The matched filter output extends for a time T on either side of the peak response. The unwanted, but unavoidable, portions of the output waveform other than the compressed pulse are known as *time sidelobes*. When the selection of the phases is made at random, the expected maximum (power) sidelobe is about $2/N$ below the peak of the compressed pulse.

Barker Codes Some random selections of the $0, \pi$ phases are better than others (where *better* means a lower maximum sidelobe level). Completely random selection of the phases, therefore, is not a good idea if compressed waveforms with low time-sidelobes are desired—and they usually are. One criterion for selecting the subpulse phases is that all the

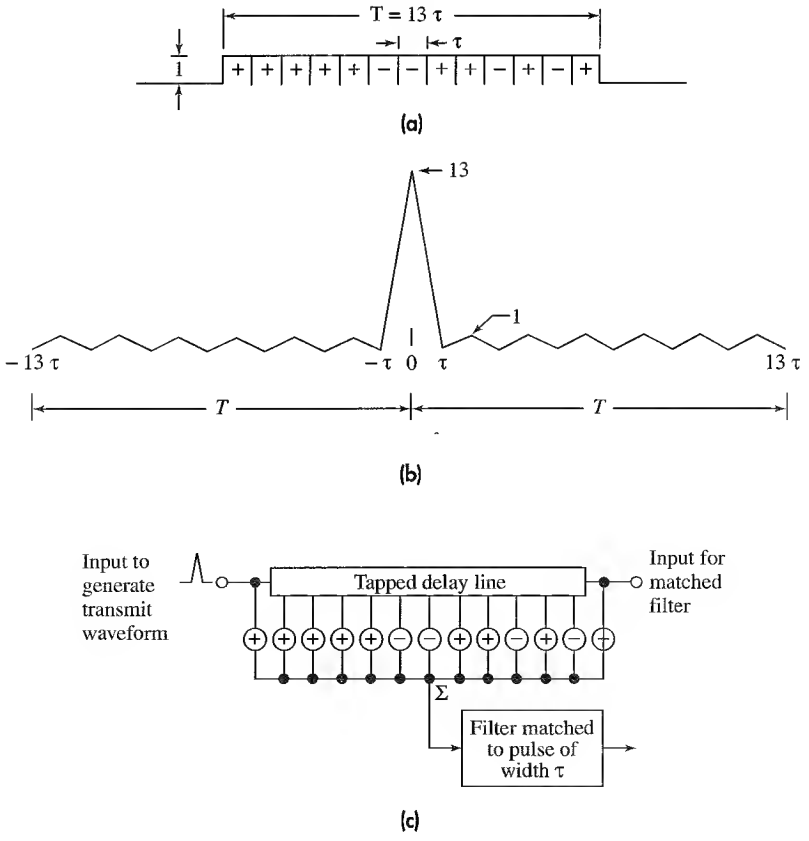


Figure 6.17 (a) Barker code of length 13; a long pulse with 13 equal subdivisions whose individual phases are either 0° (+) or 180° (−). (b) Autocorrelation function of (a), which represents the output of the matched filter. (c) Tapped delay line for generating the Barker code of length 13.

Table 6.4 Barker codes

Code length	Code elements	Sidelobe level, dB
2	+−, ++	−6.0
3	++−	−9.5
4	++−+, +++−	−12.0
5	+++−+	−14.0
7	+++−−+−	−16.9
11	+++−−−+−−+−	−20.8
13	+++++−−++−+−+	−22.3

time-sidelobes of the compressed pulse should be equal. (The reasoning here is that one can allow the small sidelobes to increase if it results in a lowering of the high sidelobes, and having all sidelobes equal is an indication that this has happened.) The $0,\pi$ binary phase-codes that result in equal time-sidelobes are called *Barker codes.* The Barker code of length $N = 13$ is shown in Fig. 6.17a. The (+) indicates 0 phase and (−) indicates π radian phase. Its autocorrelation function, which is the output of the matched filter, is shown in (b). There are six equal time-sidelobes to either side of the peak, each at a level 22.3 dB below the peak. (The sidelobe level of the Barker codes is $1/N^2$ that of the peak signal.) In (c) is shown schematically a tapped delay line that generates the Barker code of length 13 when the input is from the left. The same tapped delay-line filter can be used as the receiver matched filter if the received signal is applied from the right. The Barker codes are listed in Table 6.4. There are none greater than length 13; hence, the greatest pulse-compression ratio for a Barker code is 13. This is a relatively low value for pulse-compression applications.

Linear Recursive Sequences, or Shift-Register Codes When a pulse compression ratio larger than 13 is required, some other criterion for selecting the $0,\pi$ phases is needed. One method for obtaining a set of random-like phase codes is to employ a shift register with feedback and modulo 2 addition that generates a pseudorandom sequence of zeros and ones of length $2^n - 1$, where n is the number of stages in the shift register.[42–44] An n-stage shift register consists of n consecutive two-state memory units controlled by a single clock. The two states considered here are 0 and 1. At each clock pulse, the state of each stage is shifted to the next stage in line. Figure 6.18 shows a seven-stage shift register used to generate a pseudorandom sequence of zeros and ones of length 127. In this particular case, feedback is obtained by combining the output of the 6th and 7th stages

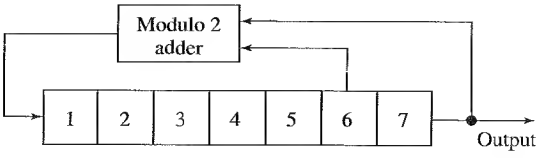


Figure 6.18 Seven-bit shift register for generating a pseudorandom linear recursive sequence of length 127.

in a modulo-2 adder. (In a modulo-two adder the output is *zero* when both inputs are the same [(0,0) or (1,1)] and the output is *one* when they are not the same. It is equivalent to base-two addition with only the least-significant bit carried forward.) An n-stage binary device has a total of 2^n different possible states. The shift register cannot, however, employ the state in which all stages are zero since it would produce all zeros thereafter. Thus an n-stage shift register can generate a binary sequence of length no greater than $2^n - 1$ before repeating. The actual sequence obtained depends on both the feedback connection and the initial loading of the shift register. When the output sequence of an n-stage shift register is of period $2^n - 1$, it is called a *maximal length sequence*, or *m-sequence*.

This type of waveform is also known as a *linear recursive sequence* (LRS), *pseudo-random sequence*, *pseudonoise* (PN) sequence, or *binary shift-register sequence*. They are linear since they obey the superposition theorem. When applied to phase-coded pulse compression, the *zeros* correspond to zero phase of the subpulse and the *ones* correspond to π radians phase.

There can be more than one maximal length sequence, depending on the feedback connection. For example, 18 different maximal length sequences, each of length 127, can be obtained with a seven-stage shift register by using different feedback connections. With the proper code, the highest (power) sidelobe can be about $1/2N$ that of the maximum compressed-pulse power. A 24-dB sidelobe can be available with a sequence of length 127. Not all maximal length codes, however, have this low a value of peak sidelobe. For example,[45] with $N = 127$, the highest sidelobe of the various maximal length sequences can vary from 18 to 24 dB below the peak. It is generally said that the more usual maximum sidelobe of a "typical" maximal-length shift register sequence is approximately $1/N$ that of the peak response. In the above example with $N = 127$, this is 21 dB. As mentioned above, a completely random selection of the phases usually results in a sidelobe approximately $2/N$ below the peak; the typical maximal-length shift-register sequence might have a sidelobe of $1/N$, and the best of the maximal-length sequences might approach 1/2N. By comparison, the Barker codes have a peak sidelobe $1/N^2$ below the peak.

Sometimes the term *code* is used and at other times the term *sequence* is used to describe the phases of the individual subpulses of a phase-coded waveform. Both terms are found in the literature and are often interchangeable when discussing pulse compression, as is the practice in this section.

The shift-register codes fit several of the tests for randomness. They are called *pseudo random* since they may appear to be random, but they are actually deterministic once the shift-register length and feedback connections are known. The fact that a pulse compression sequence is random or pseudorandom does not mean it will produce the lowest time-sidelobes at the output of the matched filter. For instance, the Barker code of length 13 in Table 6.4 is a good sequence (for its length) in that it produces a -22.3-dB peak sidelobe, but it is not what is usually thought of as "random." It does not satisfy the *balance* property of a random sequence (the number of ones differs from the number of zeros by at most one); nor does it satisfy the *run* property (among the number of runs of ones and zeros in each sequence, one half are of length two, one quarter of length three, and so forth); nor does it satisfy the *correlation* property (when the sequence is compared term by term with any cyclic shift of itself, the number of agreements differs from the number

of disagreements by at most one). Thus the Barker codes are not random in the above sense, but they produce the lowest sidelobes for their length. It should be no surprise, therefore, to find that there are better binary sequences than the shift-register sequences.

Other Binary Sequences Computer search has shown that the longest code with sidelobe level of 2 is of length 28;[46,47] the longest code with sidelobe level 3 is 51;[48] and the longest codes for levels 4 and 5 are 69 and 88 respectively.[48] It should be noted that the above sidelobe levels are almost 25 dB for code lengths varying from 51 to 88. These sidelobe levels are better than the $1/2N$ values of the best maximal-length sequences.

Doppler Effects The binary phase-coded waveform produces a thumbtack ambiguity diagram so that a bank of matched filters (with each filter tuned to a different doppler frequency) will be needed when the doppler shift of the target echo signal is large. There is, however, a more serious problem with these waveforms when there are large doppler shifts. The sidelobes in the plateau region of the thumbtack can be relatively high and a single target can result in responses from more than one filter to cause ambiguities and/or false target reports at incorrect doppler shifts. Thus the binary phase-coded waveforms might not be suitable when there is a significant doppler shift.

Quadriphase Code[49] Binary phase (biphase) coded signals using rectangular subpulses can result in poor fall-off of the radiated spectrum, mismatch loss in the receiver pulse-compression filter, and loss due to range sampling when the pulse compression is digital. These effects can be reduced by modifying the biphase codes to produce what is called a *quadriphase code.*

To generate the quadriphase code, one first starts with a good biphase code. The binary-code phases, designated as W_k for each subpulse, are transformed to the quadriphase codes, designated V_k, by means of the following transformation:

$$V_k = j^{s(k-1)}\, W_k \qquad [6.47]$$

where s is fixed and is either $+1$ or -1. Since the phases of the subpulses of a binary code are either 0 or π radians, the above equation shows that the phases of the quadriphase subpulses will be either 0, $\pi/2$, π, or $3\pi/2$ radians. Between subpulses the phase change will be either $+\pi/2$ or $-\pi/2$ radians. Each subpulse of the quadriphase code has a half-cosine shape rather than a rectangular shape. The spectrum of a cosine subpulse decreases more rapidly than that of a rectangular shape, and is therefore less likely to cause interference. The subpulse width τ is measured from the half-power points of the half-cosine. (The base-width of the half-cosine is therefore 2τ.) The overlap of the cosine-shaped subpulses results in an uncompressed pulse that has constant amplitude except for the leading and trailing edges. As mentioned previously in this section, constant amplitude output is desired of high-power transmitters since variations in the amplitude of the pulse can cause loss of transmitter efficiency. The half-cosine shape of the subpulses and the crossover at the half-power points result in constant amplitude and also eliminate phase transients that can cause spectral splatter.

The maximum value of the compressed biphase waveform (the autocorrelation function) is preserved when converted to the quadriphase compressed pulse. The peak

sidelobe level of the quadriphase code, however, can be larger than that of the biphase signal from which it was derived. The increase approaches 1.5 dB when the biphase sidelobe is very large. If, however, the amplitude of the peak sidelobe of the compressed pulse of the biphase code is unity, as with Barker codes, then according to Taylor and Blinchikoff[49] there is no increase in the sidelobe amplitude when transforming to the quadriphase code. When the peak sidelobe of the compressed pulse from a biphase code is two, the peak sidelobe of the compressed quadriphase code is increased by 0.41 dB.

If digital processing is employed, a loss in signal-to-noise ratio occurs when the sampling straddles the peak of the compressed pulse rather than being exactly at the peak. When losses due to random sampling are averaged over all possible locations within the pulse, the range-straddling loss is about 2.3 dB for binary phase codes, but is only 0.8 dB for quadriphase codes.[49] (In both the binary phase and the quadriphase codes, it is assumed in the above that the spacing between samples is the same as the subpulse spacing τ, but can occur anywhere within the subpulse. If the binary phase codes are double sampled, however, the 2.3-dB loss is reduced to 0.8 dB.) The doppler behavior of quadriphase codes is said by Taylor and Blinchikoff to be the same as the doppler behavior of biphase codes. Levanon and Freedman,[50] however, give examples where the ambiguity diagram with a nonzero doppler shift can be significantly different for a quadriphase code than for the biphase code from which it was derived. For example, the ambiguity diagram of a quadriphase code derived from a Barker code of length 13 has a diagonal ridge more like that of a linear-FM ambiguity function rather than the thumbtack ambiguity function of the Barker code. The diagonal ridge, however, is not a general feature of quadriphase codes.

Polyphase Codes The phases of the subpulses in phase-coded pulse compression need not be restricted to the two levels of 0 and π. When other than the binary phases of 0 or π, the coded pulses are called *polyphase codes.* They produce lower sidelobe levels than the binary phase codes and are tolerant to doppler frequency shifts if the doppler frequencies are not too large. An example is the Frank polyphase code[51,52] defined by an M by M matrix as shown on the left side of Table 6.5. The numbers in the matrix are each multiplied by a phase equal to $2\pi/M$ radians (or $360/M$ deg). The polyphase code starts at the upper left-hand corner of the matrix, and a sequence of length M^2 is obtained. The pulse compression ratio is $M^2 = N$, the total number of subpulses. An example of the phases for each of the 25 pulses of a Frank code of $M = 5$ is shown on the right side of Table 6.5. The basic phase increment in this example is $360/5 = 72°$. The phases are shown modulo 360°.

Frank conjectured that for large N, the highest sidelobe of a polyphase code relative to the peak of the compressed pulse is $\pi^2 N \approx 10 \times$ (pulse compression ratio). In the above example with $N = 25$, the peak sidelobe is 23.9 dB. (By comparison, the closest maximal-length shift register sequence, of length 31, has a peak sidelobe of 17.8 dB.)

Since the rate of change of phase is a frequency, examination of the matrix of Table 6.5 indicates that the frequencies of the Frank code change linearly with time in a discrete fashion. The Frank codes can be thought of as approximating a stepped linear-FM waveform. The ambiguity diagram for a polyphase code is similar to that of a linear-FM waveform, but there can be a 3- to 4-dB loss in signal at doppler frequencies that are

Table 6.5 Frank Polyphase Code

$M \times M$ Matrix Defining Frank Polyphase Code	Example of Frank Matrix with $M = 5$ and pulse compression ratio $N = M \times M = 25$
0 0 0 0... 0	0, 0, 0, 0, 0,
0 1 2 3... $N-1$	0, 72, 144, 216, 288,
0 2 4 6.... $2(N-1)$	0, 144, 288, 72, 216,
0 3 6 9.... $3(N-1)$	0, 216, 72, 288, 144,
.	0, 288, 216, 144, 72.
.	
0..... $(N-1)^2$	

The phoses of each of the M^2 subpulses are found by starting at the upper left of the matrix and reading each row in succession from left to right.

odd-integer multiples of π radians per pulse.[53] The doppler response of polyphase codes should be satisfactory with aircraft targets but might be a problem when detecting high-speed targets such as satellites and ballistic missiles.

B. Lewis and F. Kretschmer have described variants of the Frank polyphase codes, which they called P-codes.[54] They devised four P-codes that they claim to be more tolerant than the Frank code to receiver bandlimiting prior to pulse compression.

Lewis[55] has also shown that the range, or time, sidelobes of the polyphase codes can be reduced significantly after reception by following the polyphase pulse-compression network with a two-sample sliding-window subtractor for the Frank and P1 codes, and by a two-sample sliding-window adder for the P3 and P4 polyphase codes. The sliding-window subtractor is a one-sample delay whose input and output drive the subtractor. With this additional processing, the sidelobes of a polyphase code of length N are uniform and are $(2/N)^2$ relative to the peak. The sidelobe level is of unit magnitude, just as with the Barker codes. The width of the compressed pulse, however, is doubled so that the pulse compression ratio is reduced to $N/2$. There is a loss because of the doubling that Lewis estimates to be about 1 dB. For example, a polyphase waveform with an original pulse-compression ratio of 400 will have an effective pulse-compression ratio of 200 after the two-sample sliding-window subtractor or adder, and its sidelobe level will be 46 dB below the peak. This is the sidelobe level that would be achieved if there existed a Barker code with a pulse compression ratio of 200. Lewis also states that "tests of the sidelobe suppression with doppler as would be encountered in radar where the codes were useful revealed that the doppler did not significantly reduce the effect of the sidelobe suppression." There exists, however, a range-doppler coupling that is characteristic of linear FM and FM-derived polyphase codes.

Costas Codes A *frequency hopping,* or *time-frequency coded,* waveform is generated by dividing a long pulse of width T into a series of M contiguous subpulses (Fig. 6.19a). The frequency of each subpulse is selected from M contiguous frequencies within a band B (Fig. 6.19b).[56] The frequencies are separated by the reciprocal of the subpulse width (or

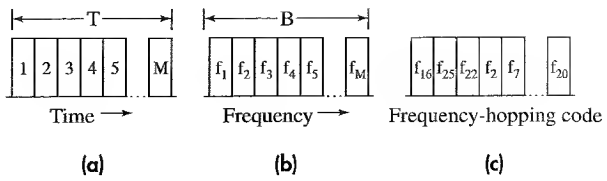


Figure 6.19 Discrete frequency-coded pulse-compression waveform. (a) Long pulse divided into M subpulses; (b) M contiguous uniformly increasing frequencies covering the band B; (c) a frequency-hopping code.

$\Delta B = M/T$); there are B/M different frequencies to choose for the subpulses; and the width of each subpulse is T/M. If the frequencies of the subpulses were to be selected so as to be monotonically increasing (or decreasing) in frequency from subpulse to subpulse, it would be a stepped-frequency waveform that approximates the linear-FM waveform, especially if the frequency and time steps are small. Its ambiguity diagram will be a ridge, like that of the linear FM. When the frequencies are selected at random, as in Fig. 6.19c, the result is a thumbtack ambiguity diagram. The pulse compression ratio is $BT = (M\ \Delta B)T = M(M/T)T = M^2$. Only $M = \sqrt{BT}$ subpulses are needed instead of the BT subpulses required for binary phase-coded pulses.

Some random selections, however, are better than others for producing an ambiguity diagram with low sidelobes; so it is not wise to choose the frequencies haphazardly. With M choices of frequencies for M subpulses, there are $M!$ different sequences. An exhaustive blind search for the best sequences is not practical except for very small values of M.

J. P. Costas has suggested a procedure for selecting the order of frequencies so as to provide well-controlled range and doppler sidelobes.[57] Costas codes attempt to have a sidelobe no greater than one unit high, so that the maximum (voltage) sidelobe level of the thumbtack ambiguity diagram is $1/M$th the central peak (voltage), where M is the number of subpulses. The sidelobes of the ambiguity diagram (in terms of relative power) are about $(1/M)^2$ relative to the central peak in those regions of the ambiguity diagram away from the central peak, and can be close to $(2/M)^2$ near the central peak.

The waveforms based on the criterion of Costas are described by an array of M rows representing frequency and M columns representing time intervals (subpulses). There is exactly one mark in each row and each column. Such an array is called an $M \times M$ Costas array if the *coincidence function* satisfies $C(r,s) \leq 1$ for all integer pairs $(r,s) \neq (0,0)$ with $|r| \leq M - 1$, $|s| \leq M - 1$. The coincidence function gives the number of coincidences of marks between the original array and its translation along the time and/or frequency axis. It may be regarded as a discrete version of the (unnormalized) ambiguity function. The parameters r and s define the amount of the translation: r is the number of integer shifts to the right or left (translated by columns), and s is the number of integer shifts up or down (translated by rows).

Consider, for example, the 6×6 Costas array of Fig. 6.20.[58] The dashed lines in this figure indicate the translation of the array two time intervals to the right and three frequency intervals up. The Xs locate the frequencies transmitted in the original array, and the Os the same frequencies in the shifted array. From this discrete representation, one point on the ambiguity diagram can be obtained. In the example of Fig. 6.20 there is only one cell in the original array and the shifted array where the X and O marks (signals) coincide. Hence the value for this point (2,3) on the ambiguity diagram is 1. The complete

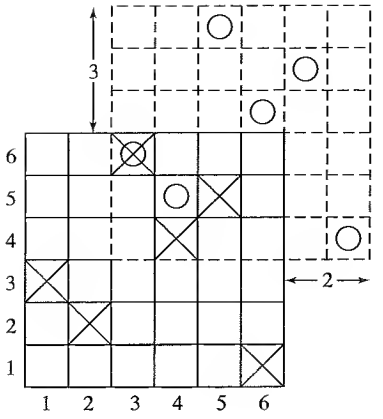


Figure 6.20 A 6×6 Costos orray indicoted by Xs along with its translation two to the right ond three up (2,3) shown doshed and indicoted by Os. A single coincidence occurs in cell (3,6). (After Chang and Scarbrough.[58] copyright 1989 IEEE.)

ambiguity diagram based on the discrete array is obtained by shifting the two arrays through all values of r and s.

A Costas code is therefore a frequency-hopped signal where there is no more than one coincidence between the original and the translated array. With no translation, there will be M coincidences, which is the peak at the center of the ambiguity diagram. Thus the maximum sidelobe (voltage) ratio is $1/M$. This is only approximate since discrete values are used. A more exact calculation of the ambiguity function, therefore, might have sidelobe levels that can be greater than $1/M$ in the vicinity of the central peak.

Golomb and Taylor[59] have analyzed the Costas codes and conjecture that $M \times M$ Costas arrays exist for every positive integer M. They show that Costas arrays exist when $M = p - 1$, $M = q - 2$, $M = q - 3$, and sometimes when $M = q - 4$ and $M = q - 5$, where p is a prime number and q is any power of a prime number. If lines were to be drawn connecting pairs of marks in the distinct cells of a Costas array, no two of these lines would be equal in both length and slope. Golomb and Taylor list the known Costas arrays for 271 values of M up to 360. They also indicate that there can be a large number of different Costas arrays for a given value of M. If $C(M)$ represents the total number of $M \times M$ Costas arrays, they give $C(M)$ for values from $M = 1$ to 13. For example, when $M = 7$ (pulse-compression ratio of 49) there are 200 different 7×7 arrays that meet the Costas criterion of only one coincidence in the discrete ambiguity diagram. For $M = 13$, there are 12,828 Costas arrays. The probability that a randomly chosen $M \times M$ permutation matrix will be a Costas array is $C(M)/M!$. Golomb and Taylor state that this probability is less than 10^{-21} when $M = 32$. Thus trial and error search is not practical for large M. Figure 6.21 is an example of a Costas array for $M = 24$ and pulse-compression ratio of 576. For this example, the maximum sidelobe in regions of the ambiguity diagram is predicted to be about 27.6 dB below the central peak when not too close to the central peak, and a bit higher in the vicinity of the central peak.

Nonlinear FM Pulse Compression[60,61] Nonlinear FM offers the advantage over linear FM of producing low time-sidelobes using a constant-amplitude waveform and a theoretically lossless matched filter. It does not experience the loss in signal-to-noise ratio associated

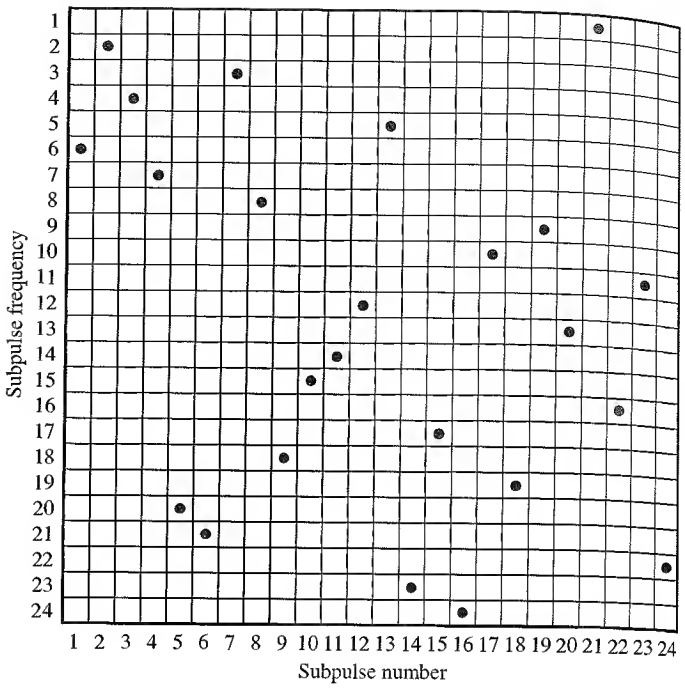


Figure 6.21 Example of frequency-time sequences for a Costas code with $M = 24$, which gives a pulse compression ratio of 576.
(After Golomb and Taylor.[59])

with the mismatched filter used to reduce sidelobes in a linear-FM pulse-compression system. A constant-amplitude envelope allows efficient generation of high power. The nonlinear rate of change of frequency performs the same role as amplitude weighting of the spectrum. If less time is spent over some part of the spectrum, it is equivalent to reduced amplitude of the spectrum. In addition, there is no significant widening of the compressed pulse. When a symmetrical nonlinear FM is used, the ambiguity diagram is that of a thumbtack; that is, it has a single peak rather than a ridge. (A symmetrical waveform is one where the frequency increases during the first half of the pulse and decreases in a similar manner during the second half of the pulse, or vice versa.) Hence, symmetrical nonlinear FM is more sensitive to large doppler shifts and is not doppler tolerant. A nonsymmetrical waveform utilizes only one half of the symmetrical waveform and has some of the range-doppler coupling characteristic of the linear FM.

Nonlinear FM waveforms result in more system complexity than linear FM. The surface acoustic wave (SAW) dispersive delay line and digital methods can be used to generate and process nonlinear FM. Nonlinear FM waveforms are designed to produce the equivalent of the classical amplitude-weighting functions mentioned previously. Examples of nonlinear FM based on 40-dB Taylor, Hamming, truncated gaussian, and cosine-squared-on-a-pedestal weightings that give low sidelobes (35 to 40 dB, or better) can be found in the literature.[61,62]

Doppler-Tolerant Pulse-Compression Waveforms A *doppler-tolerant waveform* is one whose signal-to-noise ratio out of its matched filter is relatively independent of the doppler

frequency shift (over a wide range of doppler shifts). Thus a single matched filter can be used rather than multiple filters (as in a filter bank). Such waveforms have also been called *doppler invariant.* Several of these have already been encountered in this section.

Short Pulse and Linear-FM A short pulse is doppler tolerant, as can be seen from its ridge ambiguity diagram of Fig. 6.8b. The long pulse of Fig. 6.8a, however, is not doppler tolerant. A bank of matched filters, each tuned to a different doppler shift, is required in order to detect the received signal when its doppler is unknown. The linear-FM is also doppler tolerant, as long as the absolute value of $2(v_r/c)BT$ is not greater than unity, where v_r = radial velocity of the target, c = velocity of propagation, B = bandwidth, and T = pulse duration.[62] When this product exceeds unity the peak-signal out of the matched filter will be significantly reduced and the compressed pulse widened. Except for very high-speed targets and very large values of BT, this restriction is not a concern in most radar applications; but it can be important in sonar since the velocity of acoustic propagation is much lower than that of electromagnetic waves.[63]

Linear Period Modulation The linear-FM waveform is only an approximation to a "true" doppler-tolerant waveform that is not limited by the restriction given in the above paragraph. The doppler-tolerant waveform is[64–66]

$$s(t) = A(t) \cos\left[\frac{2\pi f_0^2 T}{B} \ln\left(1 - \frac{Bt}{f_0 T}\right)\right] \qquad [6.48]$$

The amplitude $A(t)$ represents modulation by a rectangular pulse of width T [sometimes written as rect (t/T)]. The bandwidth is B and the carrier frequency is f_0. This expression for the doppler-tolerant waveform is difficult to interpret as it stands, but if the natural log is expanded using the series $\ln(1 - x) = -(x + x^2/2 + x^3/3 + \cdots)$, for $x^2 < 1$, Eq. (6.48) becomes

$$s(t) = A(t) \cos\left(2\pi f_0 t + \frac{\pi B t^2}{T} + \frac{2\pi B^2 t^3}{3 f_0 T^2} + \cdots\right) \qquad [6.49]$$

When terms greater than the first two can be neglected, Eq. (6.49) is the same as the classical linear-FM (LFM) waveform. Thus the linear FM, or chirp, pulse-compression waveform is a practical approximation to the theoretical doppler-tolerant waveform given by Eq. (6.48).

Differentiating the argument of Eq. (6.48) with respect to time, gives the frequency of the doppler-tolerant waveform as $f_0^2 T/(f_0 T - Bt)$. Inverting to obtain the period, it can be seen that the doppler-tolerant waveform is one whose period is linearly modulated with time; that is, period = $[(1/f_0) - mt]$, where $m = B/f_0^2 T$. This doppler-tolerant waveform has been called *linear-period modulation* (LPM). Because of the logarithm in the argument of Eq. (6.48), it has also been called *logarithmic phase modulation.* It is also known as *hyperbolic frequency modulation* (HFM), since the relationship between frequency and time is hyperbolic. Both the LFM and the LPM waveforms are doppler tolerant in the sense that the signal-to-noise ratio and the compressed pulse width are not significantly changed by a doppler frequency shift. However, with both waveforms there is coupling

between the range and the doppler; and the peak response will be shifted in time so that a range error can result, as evident in the ambiguity diagram of Fig. 6.9 for the LFM waveform.

The time sidelobes of the LPM waveform, after passing through the matched filter, are higher than those of the LFM waveform.[66] It has been said that weighting can reduce the LPM sidelobes to the range of -20 to -30 dB.[67] The spectrum of the LPM is not uniform but falls off exponentially as the frequency increases. The amount of exponential decay depends on the ratio of the highest-to-the-lowest frequency.

LPM pulse-compression is more likely to be important for acoustic (sonar) echolocation applications than for radar. Since the velocity of sound is much smaller than the velocity of light, the acoustic LFM waveform might result in an unacceptable widening of the compressed pulse and a lowering of the output signal-to-noise ratio. Various species of bats have been found to employ acoustic echolocation with LPM waveforms. The LPM is needed rather than LFM since the bandwidth-pulsewidth products are generally high enough to cause the absolute value of $2(v_r/c)BT$ to exceed unity.[68] The bat, which depends on acoustic echolocation to fly through dark caves with impunity and to find and catch insects in flight, has evolved a remarkable pulse compression system during its many years on earth, long before radar or sonar existed.

An accelerating target causes a frequency shift of an LPM signal.[69] The LPM waveform can be made to accommodate accelerating targets if a bank of filters is employed that is matched to the frequency-shifted versions of the received LPM waveform. It has been said, however, that wideband LFM signals "may be quite capable of achieving high acceleration tolerance without the need for separate acceleration-processing channels," even though this waveform is less doppler tolerant than the wideband LPM.[70]

Other Pulse Compression Waveforms There are several other pulse-compression waveforms that have been considered for radar. They each have some interesting characteristics; but as is true of almost everything, they also have some limitations.

Nonlinear Binary Phase-Coded Sequences[71,72] Linear recursive sequences, or shift-register codes, were discussed earlier in this section. They are formed with an n-stage shift register with a feedback logic consisting of modulo-2 additions. The number of different maximal-length sequences that can be obtained with an n-stage shift register with linear feedback logic is approximately $2^n/n$. If nonlinear feedback logic is used instead, the number of maximal length sequences increases to $2^{2^{n-1}}/2^n$. This large number is of interest when many different codes are desired for such purposes as minimizing mutual interference, providing more security to the code, or making deception jamming more difficult. With a five-stage shift register, for example, only six 31-symbol maximal-length sequences can be obtained when linear feedback logic is used. With nonlinear feedback, 2048 different 32-symbol pseudorandom sequences are available. The length of a nonlinear sequence from an n-stage shift register can be 2^n rather than the $2^n - 1$ for linear sequences.

Complementary Codes It is possible to find pairs of equal-length phase-coded pulses in which the sidelobes of the autocorrelation function of one are the negative of the other. If the autocorrelation functions from the outputs of the two matched filters are added, the algebraic sum of the sidelobes will be zero and the main response will be $2N$, where N

is the number of elements in each of the two codes. These are called *complementary codes,* or *Golay codes* after the person who first reported their existence and described how to construct them.[73] Theoretically, there are no sidelobes on the time axis when complementary codes are employed. Complementary codes can be obtained with either binary or polyphase sequences.[74]

There are two problems, however, that limit the use of complementary codes.[75] The first is that the two codes have to be transmitted on two separate pulses, detected separately, and then subtracted. Any movement of the target or instability in the system that occurs during the time between the two pulses can result in incomplete cancellation of the sidelobes. Transmitting the two codes simultaneously at two different frequencies does not solve the problem since the target response can vary with frequency. The second problem is that the sidelobes are not zero after cancellation when there is a doppler frequency shift so that the ambiguity diagram will contain other regions with high sidelobes. Thus this method of obtaining zero sidelobes has serious practical difficulties and is not as attractive as it might seem at first glance.

Welti Codes[76] These are related to Golay complementary codes in that they are used in pairs that are subtracted from one another to obtain autocorrelation functions with zero time sidelobes. They use four phases (0,90,180,270°) rather than the two phases (0, 180) of the Golay codes. They are a class of polyphase codes but have ambiguity diagrams more like those of biphase codes than the Frank polyphase codes.[77] Welti codes can present the same problems as other complementary codes.

Huffman Codes[78] So far, every pulse-compression waveform discussed in this section is of constant amplitude across the uncompressed pulse. The signal bandwidth is increased by phase or frequency modulation rather than by amplitude modulation. The Huffman codes, on the other hand, consist of elements that vary in amplitude as well as in phase. When the doppler shift is zero, they produce autocorrelation functions with no sidelobes on the time axis except for a single unavoidable sidelobe at both ends of the compressed waveform. The level of these two end-sidelobes is a design tradeoff. In one example,[79] a Huffman code of length 64 with no doppler shift has a sidelobe at each end that is 56 dB below the peak. As with other methods for obtaining zero or low sidelobes, the volume under the ambiguity diagram must remain constant [Eq. (6.46)], which means that higher sidelobes will appear elsewhere in the doppler domain. The sidelobes also degrade if the tolerances in amplitude and phase are not maintained sufficiently high or if there are too few bits in the A/D converter.

Variants of the Barker Code The pulse compression ratio of a conventional Barker code can be increased (beyond the maximum of 13) by making each element of the Barker code itself a Barker code. For example, a Barker code of length 13 in which each element is also a length 13 Barker gives a combined code length of 169 with a pulse compression ratio of 169 and a maximum sidelobe 22.3 dB below the peak.[80] Thus the pulse compression ratio is increased by increasing the length of the uncompressed pulse, but there is no decrease in absolute sidelobe level. This is called a *combined Barker code,* but it has also been known as a *compound Barker* or a *concatenated Barker.*

The sidelobe level of a conventional Barker code can be decreased by extending the time-sidelobe region beyond $2T$, where T = the uncompressed pulse width. This method, suggested many years ago by Key, Fowle, and Haggarty,[81,82] makes use of the fact that the equal-amplitude sidelobes and the main lobe of the autocorrelation function of the Barker codes are of similar shape. Because of this geometric similarity, it is possible to suppress the sidelobes by adding properly weighted and time-shifted replicas of the matched-filter output to the original output of the matched filter. That is, the autocorrelation function of the Barker-13 code, as was shown in Fig. 6.17b, is passed through a transversal filter with delays and weighting designed to eliminate the original sidelobes. If, for example, the six sidelobes that are on each side of the main lobe of a 13-bit Barker are to be removed, there can be six weighted (reduced amplitude) replicas of the Barker code autocorrelation function applied at six different times ahead of the main lobe and six different times behind the main lobe. To accomplish this the Barker code matched filter would be followed by a 13-tap transversal filter with the proper weightings at each tap to completely eliminate (in theory) the original sidelobes. As a result, new sidelobes will appear farther out in time on either side of the original Barker-13 sidelobes. These new sidelobes will be lower than the original. The new maximum sidelobe generated because of the transversal filter will be 32.4 dB below the peak instead of the 22.3 dB of the original. The loss in signal-to-noise ratio is about 0.25 dB, which is small.

Sidelobe Reduction for Phase-Coded Pulse-Compression Signals Suggestions have been made to reduce the sidelobes of phase-coded signals by following the matched filter with a transversal filter with weights that reduce the time sidelobes, similar to that described by Key et al. in the above. The weights are selected according to one of two criteria: (1) minimize the total energy in the sidelobes or (2) minimize the peak sidelobe. The number of bits (taps) in the delay line of the transversal filter can be greater than the length (number of bits) of the phase-coded signal. When the number of bits in the transversal filter is greater than that of the matched filter, the phase-coded signal is zero-padded to the same length as the filter.

One of the first papers to discuss this method was by Ackroyd and Ghani,[83] who determined the weights of a transversal filter so as to give a compressed output signal that approximates the desired output signal in the least-squares sense. The output to be minimized is a sum that represents the "energy" of the difference between the actual sidelobe levels and the desired levels. This is sometimes known as minimizing the integrated sidelobes, or ISL.[84] Starting with a Barker-13 code, Ackroyd and Ghani show that with least-squares filters of lengths 13, 41, 53, and 66, the maximum sidelobe levels below the main peak are 24 dB, 40 dB, 50 dB, and 60 dB, respectively. It is said that when the Barker-13 matched filter is replaced with the least-squares filter, the loss in signal-to-noise ratio is 0.2 dB.

This technique has also been applied to reducing the sidelobe levels of the combined Barker code.[85] A 52-element combined Barker code, for example, can be generated by combining a 4-element Barker code with a 13-element Barker code. When passed through a 200-bit minimum square-error transversal filter, a maximum sidelobe level 40.7 dB below the peak is produced. This mismatched filter results in a computed loss of 1.86 dB.

Similarly, the weights of a transversal filter can be chosen to minimize the peak sidelobe.[86] Baden and Cohen[87] show that the peak sidelobe obtained using this criterion is from one to seven dB less than the peak sidelobe obtained using the minimization of the integrated sidelobes, depending on the particular code and filter length. Loss in signal-to-noise ratio is said to be less than one dB. Compared to the sidelobes of the original output of the matched filter, reductions of 15 dB or more "are readily achieved for reasonable filter lengths." This sidelobe-weighting technique can also be used for generating mismatched filters that produce sidelobe-free regions in the vicinity of the main lobe.

Although there are several methods described for reducing the sidelobes of the output of a matched filter for binary phase-coded signals, apparently they have not had as much application as the nonmatched weighting filter used to reduce the time sidelobes of the linear-FM pulse compression waveform.

Other Aspects of Pulse Compression

Generic Compressed Signal By way of review, the general nature of the compressed pulse produced by a matched filter is sketched in Fig. 6.22. The peak of the compressed pulse is equal to $2E$, where E is the energy of the input signal. (Note, however, that the units of the peak output is in volts, not joules.) The peak value depends only on the input-signal energy and not on the signal shape. The width of the compressed pulse is approximately $1/B$, where B = signal bandwidth, regardless of how the bandwidth is obtained. When the total duration of the uncompressed signal is T, the sidelobes of the compressed pulse extend over a time $2T$. The nature of the sidelobes is determined by the shape of the uncompressed signal, $s(t)$. One of the most important aspects of pulse compression design is the selection of $s(t)$ to minimize the sidelobe level.

The major difference between pulse compression and a short pulse of the same resolution and energy is that the short pulse has no time sidelobes, whereas the pulse compression signal has sidelobes that can be mistaken for real targets or mask the presence of small targets located within the sidelobe region that extends over a time $2T$ in extent.

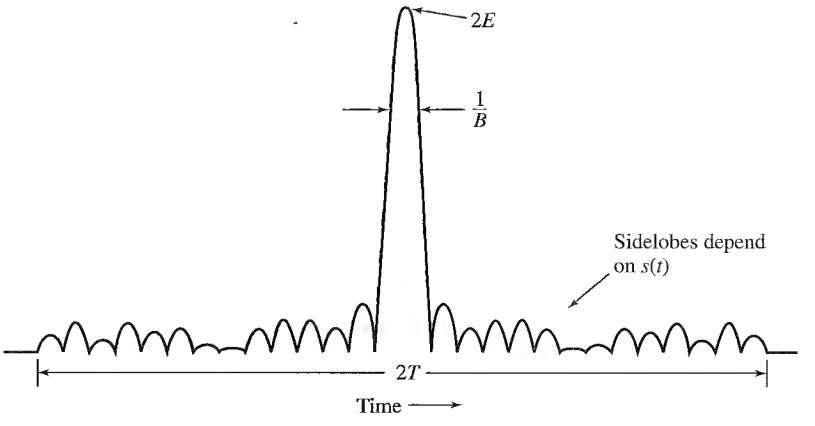


Figure 6.22 Sketch of a generic compressed pulse whose uncompressed waveform is $s(t)$, with bandwidth B, time duration T, and energy E.

Limiting in Pulse Compression Since the amplitude of uncompressed radar waveforms is constant, limiting is sometimes used in the receiver before the matched filter to insure that the input signal to the matched filter is of uniform amplitude or to provide a constant false-alarm rate using CFAR. It was said in Chap. 3 that hard limiting should not be used with MTI radar since it can significantly decrease the improvement factor. If pulse compression is used with MTI, and if hard limiting is desired for pulse compression, the limiter and the pulse compression matched filter can follow the doppler processing rather than precede it. On the other hand, when pulse compression follows the MTI, the MTI processor does not benefit from the reduction in clutter obtained with pulse compression. If there are significant instabilities in the MTI transmitter and oscillators, the use of limiters with pulse compression is different from that assumed here, as will be described later under the heading "Compatibility with Other Processing."

If MTI precedes the pulse-compression matched filter or if MTI is not used at all in the radar, then limiting can be employed with pulse compression. Even so, there are several concerns with the use of a limiter in pulse compression. These include (1) the loss incurred in the signal-to-noise ratio, (2) the suppression of nearby small target signals, and (3) the possibility of spurious (false) targets.

For phase-coded waveforms, the loss in signal-to-noise ratio decreases with the length of the code and the number of pulses integrated. Table 6.6 summarizes the loss due to limiting as given by Castella and Rudie.[88] A hard limiter is one in which the limit level is set low enough (well into the noise) so that there is no variation in the amplitude of the output signal. Phase and frequency modulations, however, are preserved. The hard limiter listed in the table corresponds to analog limiting in the IF. A digital soft limiter is defined by Castella and Rudie as one where the I and Q signal components are quantized to three bits at baseband, with two of these bits allotted to $\pm$ 2 sigma of the noise. The table indicates that one should try to have long code lengths and a large number of pulses integrated if hard limiting is used.

Table 6.6 Loss with phase-coded waveforms due to limiting[88] (probability of detection = 0.5; probability of false alarm = 10^{-6})

Type of Limiter*	Code Length	Number of Pulses Integrated	Loss in Signal-to-Noise Ratio
Hard limiter	16	1	6.4 dB
	16	8	1.8 dB
	128	1	1.4 dB
	128	8	1.1 dB
Digital soft limiter	16 to 128	1 to 8	0.4 to 0.5 dB
Digital hard limiter	16	1	8.6 dB
	16	8	3.0 dB
	128	1	2.4 dB
	128	8	2.0 dB

*See text for descriptions of limiters

The second concern mentioned above, small-signal suppression, involves two uncompressed pulses that overlap. The effect of the limiter is to cause the smaller of the two signals to be suppressed and be less detectable than if limiting were not employed. This can occur with either a linear FM or a phase-coded signal. In the overlap region the weaker signal can be suppressed by as much as 6 dB if the stronger signal has a large signal-to-noise ratio.[89] The noise is also reduced. The net result can be a reduction of the detection probability of the weaker signal, which deteriorates rapidly depending on the amount of overlap.[90]

The third concern is the possibility of generating false targets in addition to the real targets when there is overlap of the uncompressed pulses. Woerrlein[91] states that with two overlapped linear-FM signals there can be an array of false targets that appear on both sides of the two target returns. If the two signals are only partially overlapped, there will be a consequent reduction in the number of false targets. There is, on the other hand, no evidence of any false targets when two binary-phase signals overlap; but when three binary-phase signals overlap, there is a fourth signal that is false. His analysis applies only to large signal-to-noise ratios.

In the above, limiting was assumed to be intentionally included in the radar processing; however, there can be unintentional limiting that can occur when the receiver is driven to saturation by large clutter echoes or interference signals.

Cross-Correlation Properties It is sometimes desirable for more than one radar to operate simultaneously at the same frequency within the same region of coverage. Each radar can employ different pulse-compression waveforms to minimize crosstalk (when one radar accepts signals from another radar as true targets when they are really false alarms). The potential that various waveforms will cause mutual interference or false target-reports can be determined by examining the cross-correlation function (the cross-ambiguity function) between two different signals. In general, the maximum output produced in one radar by a signal from another radar with a different waveform covering the same bandwidth will usually be greater than the sidelobes of the radar's own autocorrelation function (the output of its matched filter).

If there are only two radars each using linear-FM pulse compression, one could have an up-chirp and the other a down-chirp to reduce crosstalk. Nonlinear FM waveforms can also be considered for avoiding crosstalk among various radars.

It is possible to find two Costas (frequency hopped) codes so that the sidelobe power level of the response of one in the other is no greater than $(2/M)^2$ that of the main lobe peak, where M is the number of frequencies and M^2 is the pulse compression ratio.[92] This compares with a maximum sidelobe level of $(1/M)^2$ for the ambiguity function of a Costas code, as mentioned earlier in this section. When more than two signals with Costas codes are present, the maximum cross-ambiguity sidelobes can be much greater than $(2/M)^2$.

For binary phase-coded pulses, H. Deng[93] estimated that the maximum sidelobe (power) level of the cross-correlation function for a set of k binary sequences ($k \geq 2$) is $kM/3(k-1)$ below the peak of the compressed pulse. Each of the k binary sequences is of the same length M. For $k = 2$, this expression predicts a maximum sidelobe $(M/1.5)^2$ below the peak. Thus the cross-correlation sidelobes are not as low as those of the autocorrelation function [which can be $(M)^2$ below the peak]. Using the optimization

technique known as *simulated annealing,* Deng finds sequences of codes with cross-correlation lobes close to that predicted when $k = 2$ or 3. (In his optimization he allowed the autocorrelation maximum sidelobe level to be weighted equally with the cross-correlation maximum sidelobe level.) With increasing k, the expression for estimating the cross-correlation sidelobes predicts better sidelobes than actually obtained by the use of simulated annealing, which is attributed to it being difficult to satisfy the approximations used in the estimate when k is large.

Compatibility with Other Processing Pulse compression may be used in conjunction with MTI radar, as was mentioned previously in this section. Care needs to be taken when they are used together so that if limiting is used for pulse compression it does not degrade the performance of the MTI. The increased range resolution of pulse compression can reduce the amount of clutter the radar sees so that less MTI improvement factor is required. This may be true for an ideal system; but as pointed out by Shrader and Gregers-Hansen,[94] the performance of the MTI may be no better, or even worse, than a system transmitting the same-length uncoded pulse if there are significant system instabilities that cause increased sidelobes.

The sidelobes produced in a pulse compression waveform because of system design or components that are nonlinear with frequency do not vary with time and will therefore cancel in the MTI processor just as would clutter. Instabilities due to noise from the local oscillators, transmitter power supplies, time jitter, and other transmitter noise, however, result in noiselike time-varying sidelobes that are proportional to the clutter amplitude. These noiselike time-sidelobes are not canceled by the MTI and can produce residual clutter that can cross the detection threshold and appear as targets. Shrader and Gregers-Hansen suggest a method for dealing with these noiselike sidelobes. The pulse-compression filter precedes the MTI processor, but a limiter is placed ahead of it. The limiter's dynamic range at its output is set equal to the difference between the peak transmitter power and the transmitter noise within the system bandwidth. A second limiter is placed after the pulse compression filter. The dynamic range at its output is set equal to the expected MTI improvement factor. The MTI processor follows. The two limiters cause the residue due to transmitter noise and other instabilities to be equal to the front-end thermal noise at the input to the MTI processor. The two limiters are adjustable so that when the radar is placed in the field, they can be compensated to allow for differences in clutter and the condition of the radar equipment.

Special consideration has to be given to pulse compression when used in multichannel radars, such as the three receiving channels of a monopulse tracker or the many thousands of channels of an active-aperture phased array radar. Each channel has to be highly matched (almost identical) to all the rest. The reproducibility and precision of SAW devices makes them attractive for applications where precision multiple pulse-compression units are required.[95]

Spread Spectrum Spread spectrum communication systems[96] employ coded waveforms similar to those used in pulse compression radar. The purpose of coded waveforms in communications is different, however, from their purpose in radar pulse compression. Spread spectrum communication waveforms allow multiple simultaneous use of the same frequency band by coding each transmitted signal differently from the others. In military

communications, spread spectrum waveforms also have the capability of rejecting interference as well as reducing the probability of being detected by a hostile electronic-warfare receiver. Sometimes pulse compression radars that use waveforms similar to those of spread spectrum communications have been called *spread spectrum radars*. This terminology, however, is misleading since the purpose of coded waveforms in pulse compression is entirely different from their purpose in spread spectrum communications. It is suggested that the use of the term *spread spectrum* for describing a pulse compression radar be avoided.

Comparison of Pulse Compression Waveforms As seen in this section, there are a number of different pulse compression waveforms with different advantages and limitations. Table 6.7 compares the theoretical sidelobe levels that might be achieved with various

Table 6.7 Maximum sidelobe level for various pulse compression waveforms. (Maximum sidelobe is in dB down from the peak of the compressed signal.)

Pulse Compression Ratio	Pseudorandom Sequences	Computer-Search Binary Phase	Polyphase[1]	Costas[2]
13		22.3		
15	14.0			
16			21.2	12.0
25			23.9	14.0
28		22.9		
31	17.8			
49			26.8	16.9
63	20.4			
64			28.0	18.0
73		25.2		
88		24.7		
100			29.9	20.0
112		25.4		
121			30.8	20.8
127	24			
129		25.3		
144			31.5	21.6
255	25.9			
256			34.0	24.1

[1]For polyphase codes, the maximum sidelobe is taken here to be $\pi^2 N$ down from the central peak, where $N = M \times M$ = pulse compression ratio, and M is the dimension of the matrix.

[2]For Costas codes the maximum sidelobe away from the central peak is taken to be N down from the peak, where $N = M \times M$ = pulse compression ratio and M = number of subpulses (equal to the number of frequencies). Near the central compressed peak, the sidelobes can be larger than indicated in the table.

pulse compression waveforms. The polyphase codes have the lowest predicted sidelobes in this table; but the greater the pulse compression ratio, the smaller the phase increment and the greater must be the precision. A pulse compression ratio of 900, for instance, requires a phase increment of $360/\sqrt{900} = 12°$, and a phase tolerance that is a small fraction of this increment.

Linear FM has been the waveform used most in the past in radar pulse compression. It is less complex than some others, especially if the application permits the use of Stretch. It usually requires weighting on receive to reduce the -13.2 dB sidelobes to the order of -30 dB, with a loss of about one dB. The range-doppler coupling that causes an error in the range measurement when there is a doppler frequency shift is sometimes of little consequence. If not, the true range can be obtained by use of both an up-chirp and a down-chirp. The ridge ambiguity diagram of the linear-FM waveform means that it is doppler tolerant and that a single filter can be used when there is a large doppler shift.

The linear-period waveform is related to the linear FM and is in theory a true doppler-tolerant waveform. In most radar applications, it does not seem necessary to use linear period instead of linear FM.

The nonlinear FM waveform might be more complex than linear FM, but it can give low sidelobes without the loss caused by a mismatched weighting filter. Its thumbtack ambiguity diagram means that a bank of matched filters is needed if there are large doppler shifts, further complicating the processing. In long-range radars where it is important to minimize loss, the nonlinear FM might be considered when low sidelobes are needed.

Binary phase-coded pseudorandom waveforms were sometimes considered in the past for military pulse compression radar when it was originally believed they could provide some degree of security from deception jamming or spoofing. Shift-register codes might appear random, but by examining only a portion of the code, the rest of the code can be readily predicted. Geffe[97] pointed out that the connections of an n-stage shift register that generates binary coded waveforms can be determined by elementary methods from a knowledge of $2n - 1$ successive digits of the shift-register sequence. Thus they have no inherent security. Nonlinear shift-register codes have many more options than linear shift-register codes and for this reason might be a little better for security purposes. Truly random codes might not have the limitation of pseudorandom codes; but even if these codes were fully crypto secure, they would not possess as low sidelobes as other pulse compression waveforms. Processing of binary phase-coded signals is more complex than for linear FM, and they require a filter bank to be employed when the doppler shifts are large.

A brief comparison of the linear FM and the binary phase-coded pulse compression waveforms is given in Table 6.8.

Polyphase codes have lower sidelobes than binary phase codes. They are not very doppler tolerant for large doppler-frequency shifts, but appear to be suitable for detection of targets with aircraft velocities. They could be of interest for pulse compression applications, but have not been widely used. The sliding-window modification suggested by B. Lewis[55] appears to provide significantly lower sidelobes than any other pulse compression method but with a small loss in signal-to-noise ratio.

Costas (frequency-hopping) codes achieve a partieular pulse compression ratio with fewer subpulses than phase-coded waveforms. Their sidelobes appear to be almost the same as ordinary binary phase-coded waveforms. Many more different Costas codes of a

Table 6.8 Comporison of Lineor Frequency Modulation ond Binary Phose-Coded Pulse Compression Woveforms

Property	Linear FM	Binary Phase-Coded Pulse
Time sidelobes	Good (~30 dB) when weighting on receive, and when a loss of about 1 dB can be tolerated	Can be equal to $1/2N$, and are not easy to improve; poor doppler sidelobes
Doppler	Doppler tolerant	Requires filter bank
Ambiguity diagram	Ridge	Thumbtack (but with high sidelobes in plateau)
Pulse compression filter	Single filter can be used for transmit and receive; usually analog for high resolution	Single filter can be used for transmit and receive, but with input at opposite end; usually digital
Complexity	Less complex, especially if Stretch can be used	More complex, (requires filter bank)
Application	High resolution (wide bandwidth)	Long pulses
Other	Range-doppler coupling; has been more widely used than other pulse compression	Bandwidth limited by availability of A/D converter; erroneously thought to be less susceptible to ECM spoofing

given length are available than can be obtained with binary phase codes. This property might be of interest in military radars concerned with operating against some forms of electronic countermeasures.

Complementary codes and Hoffman codes that are supposed to produce zero sidelobes along the zero-doppler time axis have interesting theoretical properties, but have serious practical limitations that make their use in radar less likely.

In engineering design when there is more than one possible method for accomplishing some desired objective, there is seldom one solution that is best for all applications. This applies as well to pulse compression. The radar designer should keep an open mind and examine the options carefully to determine the type of pulse compression waveform to be used for any particular radar application.

6.6 TARGET RECOGNITION

Early radars were "blob" detectors in that they detected the presence of a target and gave its location in range and angle, but could not provide much else about the type of target being detected. As radar resolution in range and cross-range improved over the years, it became possible to resolve the individual scattering centers of a target and infer

something about its nature. Radar began to be more than a blob detector and could provide recognition of one type of target from another.

Even without high resolution, radar has been able to recognize the general nature of a target, or scattering object, based on such information as its behavior in space and time. The frequency dependence of the cross section can also be a useful discriminant in some cases.

The various degrees of target recognition are listed below in increasing order of information required for a decision:

- *General nature of target:* Recognition that the echo on a radar display is that of an aircraft, ship, motor vehicle, bird, person, rain, chaff, clear-air turbulence, land clutter, sea clutter, bare mountains, forested areas, meteors, aurora, ionized media, or other natural phenomena. A trained and experienced radar operator with the right type of radar should be able to sort these broad classes of target echoes from one other.
- *Target type:* This includes recognizing a fighter aircraft from a multi-engine bomber aircraft, a cargo ship from a tanker, a tracked military vehicle from a truck, chaff rather than a ship, a buried rock instead of a mine; or a surface-to-air missile site from a dump site. Sometimes this coarse form of recognition has been called *perceptual classification.*
- *Target class:* This involves determining the particular class to which a target belongs among the many possible classes. For example, if the radar believes it is detecting an aircraft, is it an F-18, F-22, MIG-31, B-2, A-6, Rafale-2000, or something else? If the target is a ship, does it belong to the Aegis destroyer Class DDG-51, Aegis cruiser CG-47, Kara, Sovermeney, or so forth, or is the echo that of a chaff decoy? If it is a bird, is it a starling, mallard, or what else? The process of determining the class of the target is known as *target classification.*

The above definitions are not standardized, so one needs to be careful when hearing or reading such terms to make sure their meanings are clear. Unfortunately, this is not always the case in the literature of target recognition.

The ultimate form of target recognition is *target identification,* which involves determining the actual name of the target, its serial number, or its side number. Identification of a target usually requires a cooperative system; that is, the target must cooperate in some manner with the identification sensor. The target has to have some form of communication system, data link, or transponder system that allows it to identify itself on a regular basis or when asked for its identification by an interrogator. *Noncooperative target recognition* (NCTR) systems, of which radar is an example, obtain target recognition information without any cooperation from the target itself. In the case of radar NCTR, recognition is based on examining the characteristics of the radar echo signal received from the target.

Military cooperative identification systems are called IFF, or *Identification Friend or Foe.* They have also been known as CAI, or *Cooperative Aircraft Identification.* In civil air-traffic control, cooperative identification methods are called ATCRBS, or *Air-Traffic Control Radar Beacon System.* IFF and ATCRBS are examples of *question and answer systems,* in that an interrogator asks the question, "who are you?" and a transponder on the target automatically answers "I am a friend and my name is . . . ."

The ability to perform noncooperative target recognition with radar depends on the type of target. Ships, for example, are easier to sort by class than aircraft since ships contain many more individual scatterers with which to recognize one class from another. Target classification generally requires greater signal-to-noise ratio than does target detection since detection depends on the total signal energy, but target recognition depends on discerning the details of the target echo signature. The many small scatterers on a target can sometimes be more important for recognition than the few large scatterers that are more important for detection. The need for large signal-to-nose ratio to detect the small scatterers means that reliable target recognition usually occurs at a shorter range than does detection.

Noncooperative target recognition methods are mainly of interest to the military. By contrast, civilian needs for aircraft target recognition are usually satisfied by cooperative methods, such as the ATCRBS. NCTR sensors, such as radar designed for that purpose, have the task of recognizing one class of target from many others that might be present. Both NCTR and cooperative methods are jointly used for the important military function of *Combat Identification.* When the target is a spacecraft or satellite, target recognition is sometimes called *Space Object Identification,* or SOI. *Automatic Target Recognition,* or ATR, is a name that could apply equally well to NCTR or combat identification, but it appears to be used mainly to describe automatic methods used for the recognition of military land targets. In the civilian sector, the use of radar and other sensors to determine the nature of the natural environment is known as *remote sensing.* Remote sensing radars observe precipitation, atmospheric effects, wind shear, birds and insects; determine the earth's surface topography; explore planets and their moons (such as probing the surface of Venus beneath its ever-present cloud cover); and monitor ice conditions, the mean sea level, and the winds that drive the ocean surface.[98]

There are two aspects to target recognition. The first is to separate the target echo from its surroundings (such as clutter) and extract from the radar echo-signal information about the unique features of a target that can help distinguish one target from another. The second aspect is the method by which one makes the actual decision as to which class or type of target a particular radar signature belongs. The decision is usually, but not always, made automatically and is often based on classical pattern recognition methods or similar decision-making methods. For example, known target signatures can be stored in a computer memory (a library) and when an unknown signature is measured, it is compared with the library of signatures to see which it matches best. (This is an overly simplified statement. The actual algorithms for target recognition can be quite sophisticated.) Only the first aspect, that of information extraction, is considered here.

Target recognition, whether by radar, optics, or the human eye, is not 100 percent accurate. Even cooperative systems do not have accuracies approaching 100 percent.* A target recognition method must be able to recognize one class of target out of many tens of classes with an accuracy of perhaps 85 percent or better before it can be taken seriously. Good target recognition methods might have accuracies of 95 percent or better, which is

*The accuracy of target recognition or identification can be no better than the "availability" of the equipment (where availability is the fraction of the time the equipment can be operable). Thus if a particular IFF or an NCTR has an availability of, for example, 95 percent, the accuracy of target recognition or identification obtained from such an equipment can be no better than 95 percent.

one error out of 20 decisions. There have been many radar target recognition methods proposed and explored; but only a few have been able to achieve the accuracy and reliability required, especially when the total number of targets to be recognized at any one time is large. For this reason, more than one type of recognition method is generally employed in combination to increase the overall probability of a correct decision and approach 100 percent accuracy of a correct recognition decision.

When a target is moving in the presence of clutter, the clutter echoes can be filtered out by doppler processing, as in MTI (moving target indication) radar. If the target is stationary in the midst of clutter, MTI is not applicable and some other technique must be used to separate the target from its surrounding clutter. Some of the target recognition methods mentioned here can be applied for such a purpose. Detection of nonmoving targets is sometimes called *stationary target indication* (STI).

In the remainder of this section, several target recognition methods based on radar will be briefly reviewed. All targets are assumed to be in the clear; or, if clutter is present, the targets are separated by some means from the clutter.

One-Dimensional Imaging with High Range-Resolution Radar A radar with sufficiently high range-resolution can resolve the individual scattering centers of a target and provide the radial profile (the one-dimensional image) of the target. The radial profile might provide a measure of the length of a target in the range dimension, but the true physical target-length usually cannot be determined accurately in this manner. The ends of the target might not always provide large enough echoes to be detected, one of the ends might be masked by other parts of the target, or the aspect angle of the target with respect to the radar might not be known accurately. Even if the length could be measured with accuracy, it is not usually a good means for recognizing the particular class of most targets of interest.

The radial profile of a 757 aircraft obtained with an *L*-band air-surveillance radar having 1-m range resolution is shown in Fig. 6.23.[99] In the upper portion of the figure are superimposed seven time-aligned pulse-to-pulse radial profiles. The radial velocity of the target can be obtained by measuring the target movement from the first pulse to the last. The average of the seven time-aligned radial profiles is displayed in the lower part of the figure, from which a target dimension can be obtained. With a knowledge of aspect angle, a wingspan or aircraft length can then be estimated. Figure 6.23 is typical for a jet aircraft in that the individual resolved scatterers are relatively constant. A propeller aircraft, on the other hand, can have constant returns from the nose and the tail, but there can be large pulse-to-pulse fluctuations from the propellers, which make it possible to distinguish a propeller aircraft from a jet.

Figure 6.24 is a radial profile of a large naval ship using an *X*-band radar with a resolution of about 0.3 m. As can be seen, the radial profile of this target remained relatively the same when measurements were obtained a year later.

The radial profile has often been considered a potential method of aircraft target classification. A serious difficulty exists, however, in that the details of the radial profile can change with only a small change in aspect. Masking of one part of the target by another can occur. If there is more than one scattering center within a radar's resolution cell, the relative phases of each scatterer can change with a change in aspect. This causes

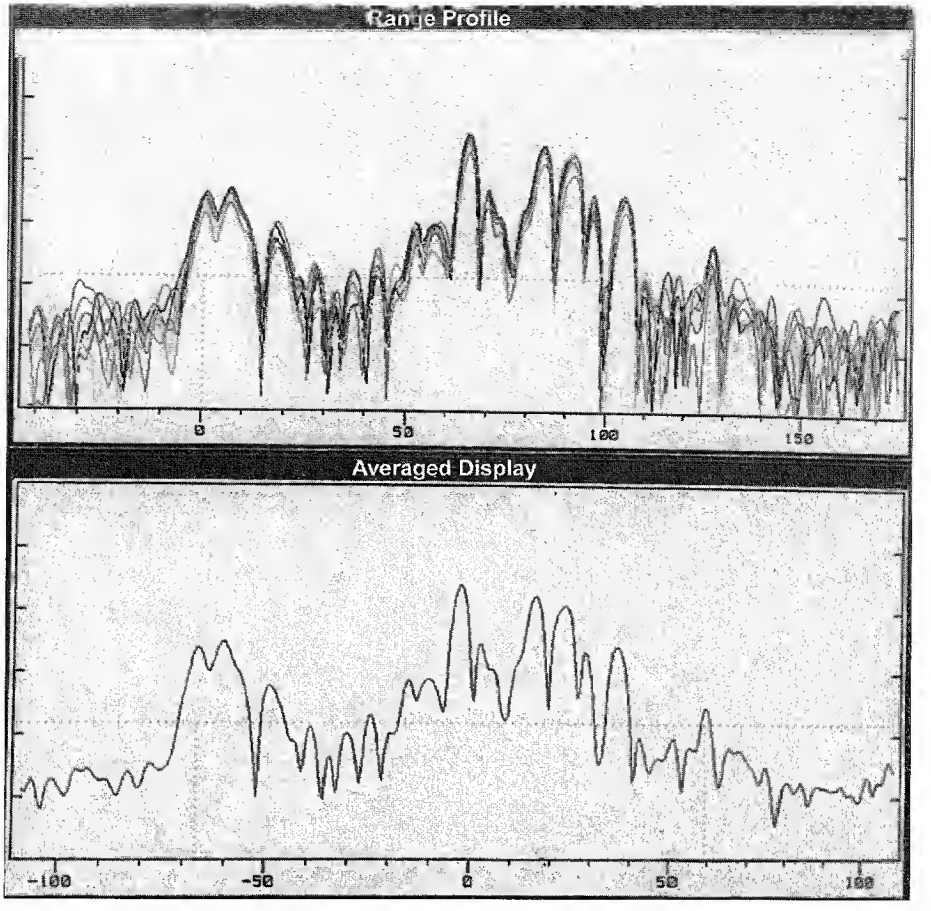


Figure 6.23 Radial profile of o commerciol 757 oircroft obtained with on L-bond rodor with o ronge resolution of obout 1 meter. Upper: the superposition of seven time-aligned pulse-to-pulse rodiol profiles. Lower: average of the seven profiles. Abscisso is in feet.
I (Provided by G. Linde of the Naval Research Laboratory.)

constructive and destructive interference and a change in the resultant cross section of the scatterers within the resolution cell. When creating a library of radial profiles to be used to match to an unknown profile, each target in the library has to be characterized by many reference profiles corresponding to different aspect angles. To make use of this library of reference profiles, the aspect angle of the unknown target must be estimated. (Observing the target's trajectory is one method for estimating the aspect.) There might need to be a large number of reference profiles stored in computer memory for each target. Furthermore, there can be many tens or even hundreds of classes of targets that need to be considered. The use of high range-resolution for target recognition starts out simple but quickly becomes complex. The result is that the use of high-resolution radial profiles for target classification is not easy to achieve in practical situations.[100,101]

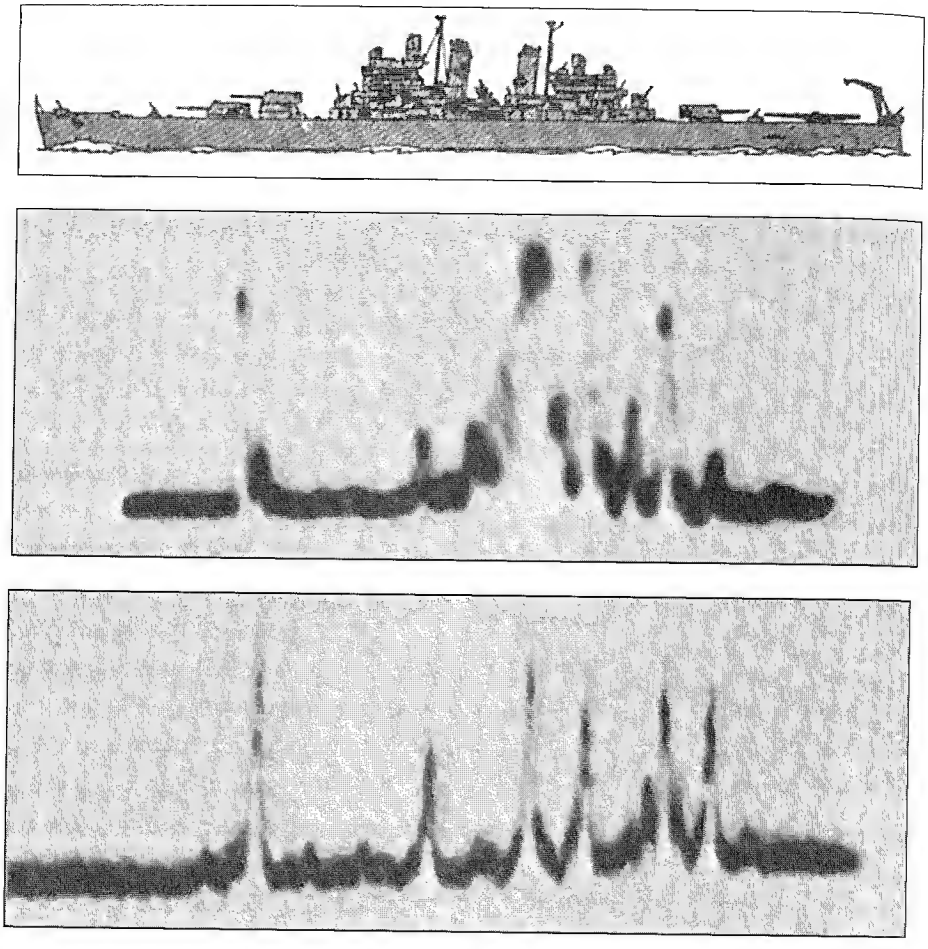


Figure 6.24 Radial profile of the former U.S. Navy gun cruiser USS Baltimore obtained with an X-band radar of 1-ft range resolution. (a) Outline of ship; (b) stern aspect; (c) bow aspect taken one year earlier than in (b).
[(Provided by I. W. Fuller, formerly of the Naval Research Laboratory)

Perceptual Classification Although it is difficult to use one-dimensional radial profiles of a target to recognize one class of target from another, it is possible to separate targets into a simpler set of general classes. This has sometimes been called *perceptual classification.* In the case of aircraft, G. Linde and C. Platis[99] were able to show from the radial profile obtained with an L-band radar, having 1-m range resolution, that targets could be separated into the following general categories: small or large jet-engine aircraft, small or large propeller aircraft, helicopters, and missiles. The perceptual classification of aircraft does not always need a dedicated special purpose radar system. It can be achieved, for example, as

demonstrated by Linde, by employing a wideband air-surveillance radar using Stretch pulse compression to obtain high range-resolution as the antenna scans by the target.

When employing high-resolution range-only information for the one-dimensional imaging of ships, it is possible to separate the ships into such categories as large ships, small ships, military ships, commercial ships, tankers, and aircraft carriers. Although perceptual classification is limited in its utility, the requirements on the radar are not as large as when determining the class of target. For reliable target classification other information usually is needed in addition to the radial profile. This added information can be the high-resolution cross-range profile, as discussed next.

Two-Dimensional Radar Imaging of Targets The two-dimensional image of a target (in range and cross-range) can be obtained by use of an imaging radar such as *synthetic aperture radar* (SAR), *inverse synthetic aperture radar* (ISAR), or more conventional high-resolution radars such as the *side-looking airborne radar* (SLAR).

Synthetic Aperture Radar SAR produces a high-resolution image of a scene of the earth's surface in both range and cross-range.[102] It can produce images of scenes at long range and in adverse weather that are not possible with infrared or optical sensors. The theoretical cross-range resolution of SAR is $\Delta x = D/2$, where D is the horizontal dimension of the SAR's real antenna aperture. SAR does not, however, accurately image moving targets. Moving targets can be seriously distorted and displaced from their true location. For instance, a railroad train, if it can be imaged at all, will be displaced from the track it is riding on. The motion of a ship target can be complicated because of its roll, pitch, and yaw. Also, all parts of the ship might not be moving rigidly together. Thus SAR is restricted to the recognition of stationary objects. One application of SAR is its military use for airborne surveillance of the battlefield and for imaging of fixed targets, as in the X-band Joint STARS. An X-band SAR image of a B-52 aircraft sitting on a runway, taken by Metratek, Inc., is shown in Fig. 6.25. Its resolution is about 1 ft. This might be compared with the ISAR image in Fig. 2.18 of a similar aircraft at VHF which has less resolution. (The resolution of published images, such as those in Figs. 6.25 and 2.18, is often degraded by the printing process and is not as good as that available from the original image that emerges from the processor.)

Inverse Synthetic Aperture Radar ISAR can be considered as a radar in which the cross-range resolution is obtained by means of high resolution in the doppler-frequency domain. Each part of a moving target can have a different relative velocity with respect to the radar, especially if there is a large rotational component of the target's motion. Resolution in doppler frequency will allow the various parts of a moving target to be resolved in the cross-range dimension. Resolution in range is obtained with either a short pulse or pulse compression so that a two-dimensional image is obtained. The cross-range resolution can be shown to be $\Delta x = \lambda/(2\ \Delta\theta)$, where $\Delta\theta$ is the change in aspect angle during the ISAR observation time and λ is the radar wavelength. Thus cross-range resolution depends on the amount of angular rotation of the target during the radar observation time. Unlike SAR, ISAR takes advantage of the target's motion to provide a two-dimensional image.

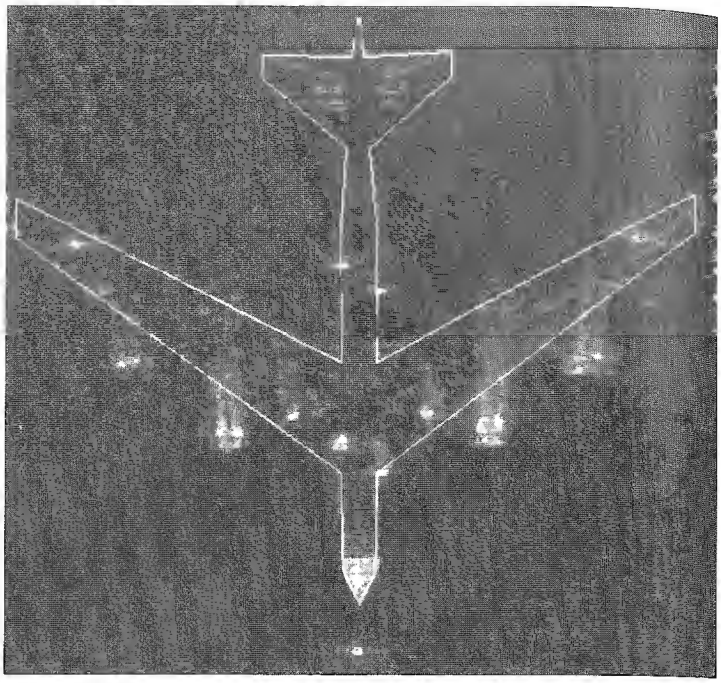


Figure 6.25 *X*-bond SAR imoge of o B-52 aircraft sitting on o runway. Resolution is obaut 1 ft. The outline af the aircroft is shown for camparisan. | (Courtesy af Ray Harris of Metratek, Inc.)

SAR and ISAR are related in that they both require a change in the aspect of a target.* In SAR, the target is assumed stationary and the radar is in motion. In ISAR, the target motion provides the changes in relative velocity that cause different doppler shifts to occur across the target. The doppler shifts from the individual scatterers are resolved by filtering. ISAR generally requires that a single large scatterer from the target be tracked to act as the reference.

Figure 6.26 is an example of an ISAR image of a commercial ship obtained with an *X*-band radar with about 2 m resolution in range and cross range.[103] The *pitch* motion of a ship causes the top of the masts to have a higher velocity than the bottom of the masts or the superstructure. These differences in velocity cause different doppler shifts. Resolution in doppler then allows the mast to be imaged. Along with conventional range resolution, the pitching motion of the ship gives a vertical profile of the target along its length dimension. *Roll* motion also provides height information, but in the plane that includes the width of the ship. Roll is not that significant for recognition since the width of a ship is small compared to its length. *Yaw* motion of the ship gives a plan view of the target. As the ship pitches, rolls, and yaws, ISAR might produce a vertical profile along the major axis of the ship, or a vertical profile along the minor axis of the ship, or a horizontal plan

*SAR and ISAR are variations af the same phenamenan. This is seen in ISAR by requiring a change in the target aspect af $\Delta\theta = \lambda/2\Delta x$. In SAR the change in fixed-target aspect is determined by the beamwidth of the real radar aperture $\Delta\theta \approx \lambda/D$. Using this expressian far beamwidth in the expressian far the SAR cross-range resalutian $\Delta x = D/2$, it is faund that $\Delta x = \lambda/2\Delta\theta$, which is the same as the ISAR crass-range resolutian.

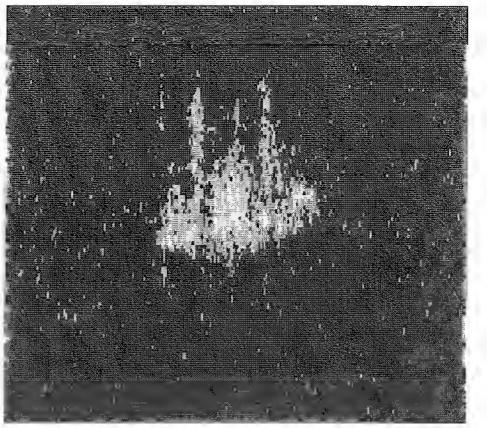


Figure 6.26 ISAR image, before image processing, of a commercial ship (17,000 ton) obtained with an *X*-band radar having 2 m resolution. The vertical scale in this image is slightly exaggerated. Note that "radar eyes" are not like "optical eyes," yet useful information can be obtained from a series of such images.
(Provided by Ronald Lipps of the Naval Research Laboratory)

view, or some combination of the three that results in the image appearing as a perspective view. Since the angular rates of pitch, roll, and yaw are not generally known, an ISAR image does not represent true distances in cross-range, as it does in range. This doesn't seriously limit, however, the ability to recognize one class of ship target from the other.

Since different ISAR images are obtained as the ship pitches, rolls, and yaws it might require many tens of seconds of observation to acquire images suitable for classification. For ship targets, relatively long observation times are not necessarily a serious problem since time is not as urgent as it would be for aircraft recognition. Sometimes an experienced operator can recognize the ship by simply viewing its ISAR image. In most cases, however, the operator has to employ a more structured technique, especially when there are a large number of possible ship classes to which a target might belong. An operator can use a combination of three different techniques: (1) *measurement of the relative locations* of the major scatterers along the bow-stern axis, such as masts, superstructure breaks, guns, and missile launchers; (2) *feature descriptions* which characterize scattering features by their degree of match to descriptive templates such as the shape of the stern (straight, curved, or rounded) or the type of masts (pole, lattice, or solid); and (3) *shape correlation* by visually comparing on the same display a single ISAR image with a superimposed wire-frame model of the candidate target which has been transformed to match the orientation of the target image produced by the ISAR.

Ships are generally large radar targets so that with a good radar the range at which ship recognition can be performed is basically limited by the horizon. It can be tens of miles with a ship radar, 200 miles with an aircraft radar, or much greater ranges with a spaceborne radar.

The classification of aircraft with ISAR is much more difficult than the ISAR classification of ships. First, an aircraft has many fewer scattering centers than does a ship. With fewer scattering centers or features presented by the unknown target, the recognition decision will be less accurate. (ISAR images of an aircraft have seldom been shown in the past without an outline of the aircraft superimposed to allow the viewer to recognize what is being seen.) Second, an aircraft does not usually experience the relatively large pitch, roll, and yaw motions of a ship, so that the cross-range resolution might not

be good enough to isolate the scattering features important for target recognition. Since resolution depends on having a sufficient change of aspect angle, the aircraft must be observed for a relatively long time if it is moving on a straight-line path. Usually a deliberate maneuver by the target aircraft is needed in order to obtain the required change in viewing aspect ($\Delta\theta$) needed for acceptable high-resolution ISAR imaging. Third, aircraft have much smaller radar cross sections than do ships. Therefore, the ranges at which aircraft recognition can be performed, even if the other two limitations are not present, will be much less than those obtained with ships. These difficulties combine to make ISAR imaging of aircraft less attractive than the ISAR imaging of ships.

Although microwave ISAR might not provide aircraft target recognition as well as it does ship recognition, it has been successfully used as a diagnostic tool for understanding the nature of radar scattering from aircraft and the design of low radar cross-section aircraft. Radar images of aircraft in flight can be obtained from both ground-based and airborne radars,[104] so long as sufficient relative motion is achieved between the radar and the target. An ISAR image of a KC 135 jet aircraft (a military version of the Boeing 707) was shown in Fig. 2.18.[105] The aircraft to be imaged flies behind the aircraft carrying the imaging radar in its tail. The radar aircraft maneuvers from one side to the other so that its radar can obtain the relative velocity (and doppler shift) needed to produce an image. Whether one calls this radar an ISAR or a SAR is immaterial. It relies on the relative motion between the imaging radar and the various scatterers that make up the aircraft to be imaged. The advantage of this high-resolution imaging method is that the individual scatterers that contribute to the backscatter echo can be readily recognized and their contributions to the total target cross section can be determined.

An innovative, experimental aircraft-recognition radar is that reported by Steinberg[106,107] to observe commercial aircraft flying into Philadelphia International Airport. It employed a ground-based ISAR to image aircraft passing at relatively close range (typically 3 km). Target aircraft were viewed by the radar in the vicinity of broadside where the rate of change of angle was large so that good cross-range resolution was obtained. A single transmitter was used with two receivers. One receiver was colocated with the transmitter, the other was separated by 25 m. Both were operated as monostatic (single-site) radars. Their physical separation allowed two different target images to be obtained. A third image was obtained by operating the two receiving antennas as an interferometer. The resulting highly processed image of an L-1011 commercial aircraft, with 47-m wingspan and 54-m length, is shown in Fig. 6.27a, along with a plan view drawing of the L-1011 (Fig. 6.27b) for comparison. Figure 6.27a is not that of a single image, as is Fig. 6.27c, but is the superposition of the three images mentioned above (one from each receiver and the two as an interferometer) as well as the inverted images of the three so as to take advantage of the known symmetry of the aircraft about the longitudinal axis. This is an exceptionally good "ISAR" image because of the short range, large signal-to-noise ratio, large change in aspect angle, diversity overlay of multiple images, lack of clutter, and inclusion of the inverted images so as to depict scatterers that might have been masked by the fuselage or the tail. For comparison, several range profiles of the same L-1011 aircraft are shown in Fig. 6.28 at different aspect angles to illustrate the dramatic change in profile with aspect.

Generally, aircraft in normal (nonmaneuvering) flight do not change their aspect angle sufficiently to make good ISAR images with *X*-band radar. Also there are few

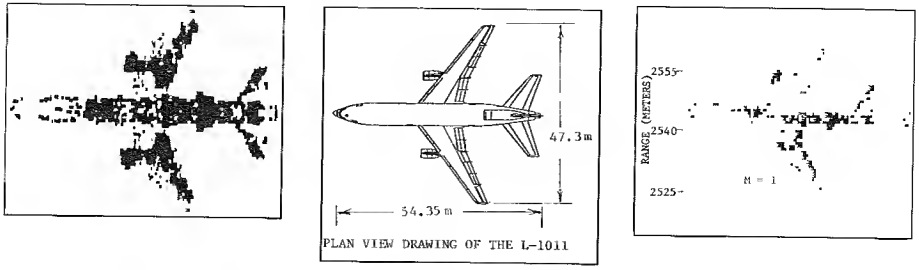
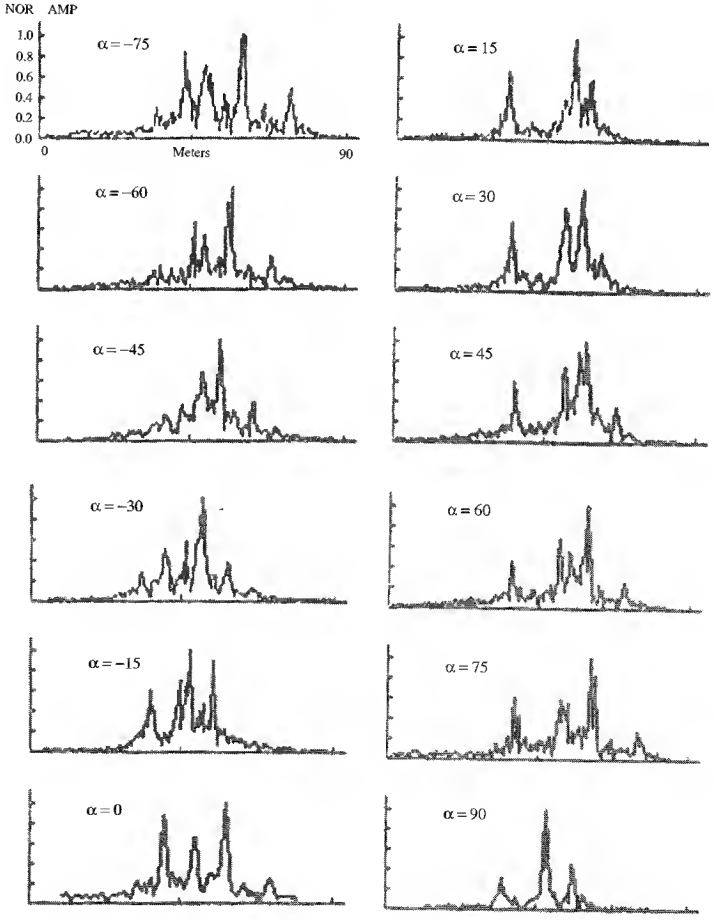


Figure 6.27 (a) ISAR radar image af an L-1011 aircraft made up by superimpasing three independent images alang with their individual inverted images (see text). (b) Outline drawing of the same aircraft shawn far camparisan. (c) One af the single images that was used as a part af (a).
(Courtesy of Prof. Bernard Steinberg of the Univ. of Pennsylvania.)

Figure 6.28 One-dimensionol profiles af the Lackheed L-1011 os a function of ospect ongle, abtoined fram Fig. 6.27o. Head-on aspect occurs ot $\alpha = 0$ deg; positive ongles view the port side of the aircraft.
(Courtesy of Prof. Bernard Steinberg of the Univ. af Pennsylvania.)



distinctive features in an aircraft image that allow it to be recognized from other similar aircraft. Thus ISAR imaging of aircraft in normal flight has been disappointing. The ISAR imaging of aircraft at *W* band (94 GHz), however, requires only one tenth the change of target aspect necessary at *X* band for the same resolution. Because scattering occurs from variations in the target surface that are comparable to the radar wavelength, it is also likely there will be more scatterers imaged at millimeter waves than at lower frequencies if the signal-to-noise ratio is high enough. In one experimental investigation,[108] it was found that, compared to ISAR images at *X* band, ISAR images of a small Piper Navajo aircraft made at 49 GHz "showed scattering from small details that tended to fill in the target shape and produce an outline view of the target." Furthermore, the target cross section at 49 GHz, when averaged over 360°, was 8 dB greater than the 9-GHz cross section. The largest increase was 19 dB, and was at nose-on incidence. Millimeter-wave ISAR, therefore, should produce better results than does ISAR at the lower frequencies.

Earth-based radars were used in the past to make ISAR images of the moon, Venus, and Mercury; but this technique has been replaced by planetary SAR radars that can orbit close to a planet and obtain superior resolution than an earth-based radar. ISAR for planetary exploration was called *delay-doppler mapping* by the radar astronomer.[109]

Radar Cross-Section Modulations Radar echoes from moving parts of a target can help recognize one type of target from another or determine the class to which a target belongs. Modulations of the radar backscatter due to propellers, helicopter rotors, jet engines, tank treads, rotating antennas, rotating machinery, the wing beat of birds and insects, and even the heartbeat of a human are examples that might be employed for target recognition. Narrowband (long pulse) as well as wideband (short pulse) waveforms can be used for extracting these modulations.

Propeller Modulation Aircraft propellers cause a distinctive modulation of the echo signal. The modulation depends on the rotation rate of the engine, the aspect angle, the number of blades on the propeller, and the shape of the propeller. An example of the doppler spectrum obtained with a coherent S-band radar is shown in Fig. 6.29 for a DC-7 aircraft,

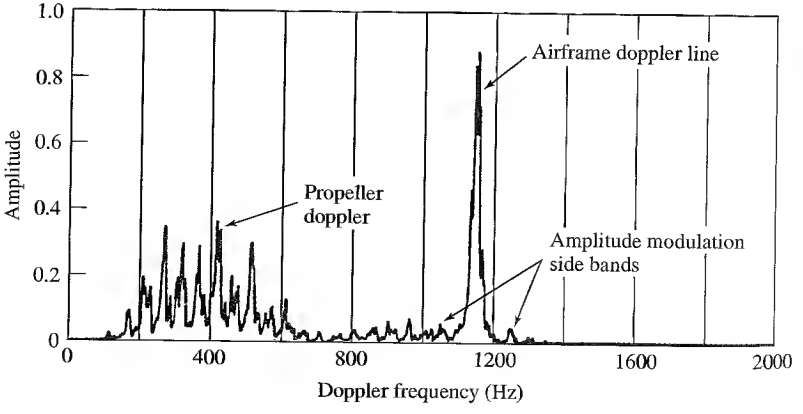


Figure 6.29 Spectrum of the propeller modulation from a DC-7 four-engine commercial aircraft taken with an *S*-band radar and one-second dwell time. (From R. Hynes and R. E. Gardner of the Naval Research Laboratory.[110,116])

a four-engine propeller-driven commercial aircraft.[110,111] The "airframe doppler-line" is the doppler frequency shift from the aircraft itself. The low-level amplitude modulation of the airframe doppler signal is caused by the propeller blade "chopping" a portion of the radar energy reflected from the airframe. At a frequency lower than the airframe line, there are spectral lines due to the radar echo from the rotating propeller blades themselves.

Propeller modulation has not been used to recognize one type of propeller aircraft from another, but its absence can be used to readily distinguish a jet-engine aircraft from a propeller-driven aircraft. In a high-resolution imaging radar, propeller modulation might locate the position of the engines, which could be helpful for target recognition.

Helicopter Blade Modulation Helicopters can be distinguished from fixed-wing aircraft by the characteristic modulation of the radar echo that comes from the rotating blades and the rotating hub structure supporting the blades.

A "blade flash" occurs every time one of the rotating blades is perpendicular to the radar beam direction.[112,113] In this position the radar cross section is a maximum. For a two-blade rotor, there are two flashes per revolution. Rotors with an even number of blades produce a number of flashes per rotation equal to the number of blades. An odd number of blades produce twice the number of flashes per revolution as the number of blades. (This can be verified by drawing a sketch and noting that there will be a separate flash from the front and rear when there are an odd number of blades, but a simultaneous flash from a front and a rear with an even number of blades.) The linear speed at the tip of the blade of helicopter is not highly dependent on the type of helicopter. It varies from about 210 to 230 m/s. If L = the length of the blade in meters, N = number of blades, and the speed of the tip is taken to be 210 m/s, the blade period in seconds for an even number of blades is

$$T_B = \frac{2\pi L}{210N} \qquad [6.50]$$

The period for an odd number of blades is one half that for an even number. Collot[112] points out that the blade period can be used to separate one type of helicopter from another since the rotation speed is constant whatever the type of flight.

Collot also states that the radar cross section, in square meters, for the blade flash is approximately aL^2/λ, where L = length of the blade, λ = wavelength, and the constant $a = 0.5$ for the front edge of the blade and 0.1 for the trailing edge. (The front edge has a higher cross section than the trailing edge since it is blunt and the trailing edge is sharp.[114]) Because of the difference between front and rear cross sections, a rotor with an odd number of blades will have alternating values for the blade flash (since the front edge and rear edge are not seen simultaneously). The echoes from the blade flash with an even number of blades, however, will be the same strength each time. The duration of the blade flash depends on the rotor speed, the length of the blade, and the radar wavelength. Assuming that the reflection from the blades can be approximated by a $(\sin x)/x$ relation and that the angular extent of the reflected echo is λ/L radians, the duration of the blade flash is approximately $t_f = \lambda/420$ for an even number of blades and twice this for an odd number (again assuming that the velocity of the blade tip is 210 m/s). At X band, with λ = 3.2 cm, the duration of the flash with an even number of blades is about 75 μs.

The intermittent nature of the short-duration helicopter blade flash might go unnoticed if the time on target is less than the time between flashes [which is the blade period of Eq. (6.50)]. In order that the helicopter blade flash be intercepted by the radar, a high prf and a long time on target are required.

The frequency spectrum of the radar echo from a helicopter has energy about the airframe doppler line due to reflections from the rotating hub that holds the blades. The helicopter spectrum will be asymmetrical since the echo from the approaching blade will be at a higher frequency than the airframe line and be of a larger amplitude than the echo from the receding blade at a lower frequency. Collot states that recognizing that an echo is from a helicopter and not a fixed wing aircraft is "immediate," and the recognition "of the type of helicopter is given in a very short time with a probability better than 95 percent." (He does not state, however, how many different types of helicopters were involved in obtaining this probability.)

In addition to the characteristic features mentioned above, the classification of helicopters might also include the echo characteristics of the tail rotor, whether there are single or twin rotors, and the configuration of the rotating hub.[115] Bullard and Dowdy[113] point out that to recognize a helicopter based on the spectral characteristics of its echo signal, the radar signal sampling must be at the Nyquist rate or higher to prevent aliasing. For *X*-band radar they state that a minimum prf of 30 kHz is required.

Jet-Engine Modulation (JEM) The radar echo from a jet-engine aircraft will be modulated by the engine's rotating compressor when looking in the vicinity of the nose. There might also be radar echo modulations from the turbine when looking from the rear of the aircraft; but these usually have a much smaller echo than the echo from the compressor.[110] The characteristic modulations of the radar echoes from aircraft can be used for recognizing one type of aircraft from another, or more correctly, one type of aircraft engine from another. Even though the jet engines are set back at a distance from their air intake and are entirely enclosed except for intake and exhaust ducts, there usually is sufficient propagation down the duct at microwave frequencies to obtain an echo signal from the compressors. Radar echoes from the compressors are sometimes obtained out to angles 60 deg from the head-on.[116]

An example of the *jet-engine modulation* (JEM) produced by a multiple-engine jet viewed head-on by an *X*-band radar is shown in Fig. 6.30.[110] Shown in this spectrum are the airframe line and the lower frequency sidebands of the compressor modulation. The component labeled c is displaced from the airframe line by a frequency $\Delta f = nb_c$, where n is the number of revolutions per second of the compressor and b_c is the number of blades on the first-stage compressor. There are other engine spectral components that are displaced from the airframe doppler line by integer multiples of Δf. The upper sidebands (not shown here) are symmetrical in frequency with the lower sidebands (with reference to the airframe line), but are lower in amplitude. Because of the high rotation speed of the engine components, aircraft jet-engine modulations are likely to be as high as ten to twenty kilohertz at *X* band. This requires high sampling frequencies (prfs) to avoid aliasing.

Mathematical analysis of jet-engine modulations indicates that the overall JEM signal can be decomposed into an amplitude modulation component and an angle (phase or frequency) modulation component that can be considered separately. The details of the line spectrum of

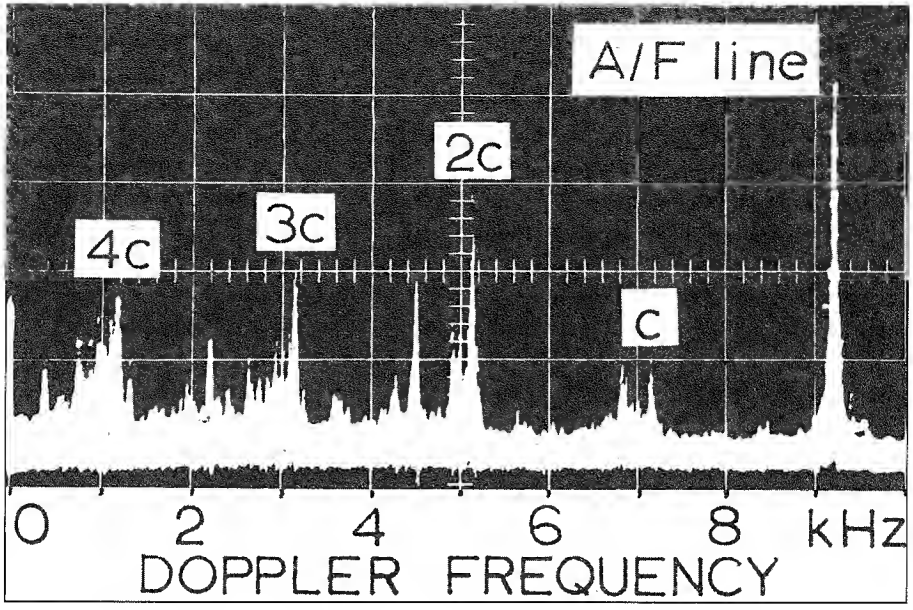


Figure 6.30 Spectrum of a multi-engine jet aircraft taken at 0 deg aspect angle showing the airframe line due to the aircraft's relative velocity and modulations from the compressor. (From R. Hynes and R. E. Gardner of the Naval Research Labaratory.[110,116])

the echo signal depend on many things including the number of blades on the first and second stages of the compressor. A minimum time of observation is needed to obtain a meaningful spectrum suitable for target recognition, which is said to be at least 25 ms.[117]

Polarization Response and Target Recognition The radar echo from a target can be affected by the polarization of the radar signal. Many research engineers have hoped that the differences in the echo from targets viewed with different polarizations might be used as a means for distinguishing one from another.[118] This is possible in simple cases, as, for example, distinguishing between a long thin rod and a sphere by using a rotating linear polarization (rotating electric field vector).

In the more general situation with realistic targets, such as aircraft, there have been many attempts to use the polarization matrix of the target echo and other polarization descriptors[119–121] for target recognition. Target recognition by means of polarization generally requires measurement of what is called the *polarization matrix.* This is obtained from the radar echo signals received on both horizontal and vertical polarization when horizontal is transmitted, and on the received horizontal and vertical polarization signal when vertical is transmitted. The polarization matrix is a 2×2 complex scattering operator (phase as well as amplitude) that characterizes a target's scattering properties. It may be expressed as

$$S = \begin{bmatrix} S_{HH} & S_{VH} \\ S_{HV} & S_{VV} \end{bmatrix} \qquad [6.51]$$

where the first subscript represents the polarization of the receiving antenna and the second subscript is the transmitted polarization, H is for horizontal polarization and V is for vertical polarization. For example, S_{HV} is the complex target backscatter signal (amplitude and phase) received with a vertically polarized antenna when a horizontally polarized signal is transmitted. The HH and VV are both *co-polar* components and the HV and VH are both *cross-polar* components. In general $HV = VH$.

Although theory might indicate that the polarization response of a target depends on the nature of the target, there has not been the desired success in applying polarimetric methods for practical target recognition. There are several reasons why this has been so: (1) Multipath signals from the surface of the ground or from clutter and structures along the propagation path can modify the polarization of the signal. (2) Practical antennas are limited in the purity of their polarization response, and there is always some finite response to the cross polarization which sometimes might be relatively large. (3) The two orthogonal polarizations (horizontal and vertical) have to be transmitted on different pulses; and if the target changes its aspect during the time between pulses, the polarization response will be distorted. (4) Multiple (unresolved) scatterers or targets within the radar resolution cell will result in a composite signal with a polarization different from any of the individual scatterers. Giuli[122] in his classic paper on polarization diversity states that low-resolution radars "intrinsically provide quite limited classification properties." He points out, however, that wideband radars that can resolve the individual scatterers of a target might make exploitation of the polarization information more profitable by avoiding the changes in echo polarization caused by unresolved scatterers.

Even though the use of polarimetry for target recognition has excited much interest, one can conclude that as a means for recognizing complex targets its effectiveness in the past has been disappointing.

Other Radar Target Recognition Methods Several other radar methods for the recognition of targets are briefly mentioned below. Most can be thought of in terms of aircraft recognition, although they are not limited to aircraft.

High Range-Resolution with Monopulse[123] The use of monopulse radar for target tracking in angle was discussed in Sec. 4.2. If the monopulse radar has sufficiently high range-resolution so as to isolate the major scattering centers of a target, then a measurement of the angular location of each scatter in azimuth and elevation can be made. Thus high range-resolution with monopulse can provide a 3D-like "image" of a target that might be used for target recognition. In addition to requiring high range-resolution, a limitation is that the range must be short enough so that angle measurements can be made with enough accuracy to obtain adequate cross-range measurements of the major scatterers.

Fluctuations of Radar Cross Section with Aspect Angle Section 2.7 indicated that small changes in the aspect angle of complex (multiple scatterer) targets can cause major changes in the radar cross section. This has been considered in the past as a means for target recognition, but it has not had much success. In some respects this method is related to the imaging of moving targets by means of ISAR, discussed previously, except that the highly

successful ISAR technique employs both phase and amplitude. Only amplitude information is available when cross-section variations are observed.

Resonance Region Response There have been several methods proposed for target recognition based on using a number of distinct frequencies in the resonance region of target scattering. (In the resonance region the dimensions of the target are comparable to the radar wavelength.) One method is derived from the ramp response of a target and develops a so-called *feature space* based on the amplitude, phase, and/or polarization of each distinct frequency that observes the target.[124,125] A related method employs the complex *natural resonances,* or *poles,* of the target to characterize the target.[126,127] The *Singularity Expansion Method* has also been applied to the target recognition method that employs frequencies in the resonance region.[128]

These methods have a good theoretical basis; they generally do not depend on knowing the target aspect angle; and they have been tested with computer simulations and experimental model measurements at scaled frequencies. They are not practical, however, for most applications of target recognition since they require frequencies in the resonance region, which for aircraft targets would be in the HF portion of the spectrum. For ships, even lower frequencies would be required. These recognition methods depend on the so-called *late-time scattering* due to creeping waves that travel around the target. Even if long radar wavelengths were no problem, it has been pointed out[129] that the effect of noise can mask the creeping waves that occur beyond the first creeping wave so that there might not be enough information to determine the complex resonances.

Another related method is to transmit a radar waveform that is based on knowing ahead of time the natural resonances of the target so that the desired target echo signal can be readily recognized from the echo signals from targets with a different set of natural resonances. One such method is called *K-pulse;* another is called *E-pulse.*[130] They require that the radar transmit a special signal for each target to be recognized. Rather than transmit special waveforms it might be better to transmit a more normal wideband radar pulse and perform target recognition by convolution processing in the receiver to recognize a particular target signal.[131]

Since they have been difficult to apply in practice, the above resonance region methods have been more of academic interest than for practical target recognition applications.

Nonlinear Scattering Effects When a metal comes into contact with another metal, oxides can form and the junction can act as a nonlinear diode. A radar echo signal reflected from such a metal-to-metal junction will contain higher harmonic components, with the third harmonic usually being the largest.[132–134] Thus if a frequency f_1 is transmitted, detection of the third harmonic, $3f_1$, indicates that the scatterer has metal-to-metal contacts. This nonlinear response has been of interest as a possible method for recognizing certain types of metallic targets. If the third harmonic of the echo signal is used, the transmitted signal must not have significant third-harmonic radiation of its own. A method for avoiding the third harmonic problem is to simultaneously transmit two frequencies f_1 and f_2, and tune the receiver to a strong cross-product such as $2f_1 \pm f_2$. The use of nonlinear target properties for radar has sometimes been called METRRA, which stands for *Metal Reradiating Radar.*

The amount of signal returned from a nonlinear contact at a harmonic frequency is a nonlinear function of the incident field strength. Thus the nonlinear target cross section depends on the peak power. The normal radar equation does not apply.[135,136] The range dependence in the radar equation might be as the sixth power or higher, instead of the usual fourth power. Because of the nonlinearity, in this specialized case high peak power is more important for detection than is high average power.

Although radar detection based on harmonics produced by nonlinear metal-to-metal contacts has some interesting attributes for applications (such as deteetion of stationary targets in clutter), it is difficult to apply since the echo signals at frequencies other than those transmitted can be very weak.

A related effect is based on the modulation of the scattered signal when metal-to-metal contacts are opened and closed so that the current distribution on a metal target is modified.[137] An example is the intermittent contacts due to the moving wheels of a train. This effect is sometimes called RADAM, which means *Radar Detection of Agitated Metals.*[138] The echo signals from RADAM are usually too weak to be of interest for satisfying the needs of target recognition.

Target Track History In military operations it is important to determine whether a target held in track by a radar is a friend or is a hostile threat. There are several military recognition or identification methods that might be used for separating friend from foe; but the radar itself might be able to assist in the recognition task based on the history of a target's track. If one knew where an aircraft or ship originated, where it has been, and what it is currently doing, it ought to be possible in many cases to determine whether the target is potentially hostile or not. A radar with a good automatic tracking system and with adequate memory is required, as well as logic to recognize the trajectories of potentially hostile targets from ones that are not.

In some cases, military aircraft recognition is accomplished by having friendly aircraft fly in pre-designated corridors in a known manner that the radar is able to observe.

Signal Intercept and Direction Finding Combined with Radar So far, this section has been concerned with radar methods for recognizing one target from another. Sometimes, however, the use of radar with other sensors can have a synergistic effect on the quality of target information. An example is the use of an electronic warfare intercept receiver (a part of ESM, or Electronic Support Measures*) and direction finder (DF) in combination with a radar. The intercept receiver might be able to recognize the type of target by the characteristic signals it radiates. The direction to the target can be obtained by the DF, but electronic warfare equipment usually cannot obtain the target's location in range. Radar, on the other hand, provides target location and track. The two together can locate, track, and obtain target recognition if the radar target track can be properly associated with the DF angle track. The angle accuracy of ESM/DF is often much poorer than the angle accuracy of radar. Nevertheless, algorithms can be developed that allow the ESM/DF information to be associated with a target under track by the radar so that target recognition provided by the ESM can be applied to the radar target under track.[139–141]

*ESM is also known as Electronic Warfare Support, or ES.

Target Recognition Applications There are many areas of application where the radar recognition of a target echo signal is important. Some have already been mentioned previously in this subsection, others are briefly described below. All of these require the intelligent extraction of information from a radar signal, a radar capability which this chapter has tried to introduce.

Military Combat Identification Studies of fratricide, which is the unintentional infliction of casualties on one's own forces, show that fratricide occured in past military hostilities at rates that typically vary from 10 to 20 percent, or greater.[142] Casualties from "friendly fire" have been considered by the military as an unwelcome consequence of war that needs to be kept to a minimum. To minimize the fratricide rate, military forces employ a number of recognition and identification methods as well as strict rules of engagement. Reliable, accurate, and secure combat identification is necessary to reduce fratricide. Both cooperative and noncooperative methods are employed in combination to increase the probability of a correct decision in a timely manner. The particular methods differ depending on whether the target is an aircraft, ship, ground vehicle, or the soldier on the ground. Radar target-recognition methods include engine modulations, track history, SAR, and ISAR.

Ballistic Missile Target Discrimination Defending against a single ballistic missile is a demanding task, but it can be done. The difficulty in ballistic missile defense escalates tremendously when the reentry vehicle carrying the warhead is accompanied by *penetration aids* that might consist of the spent booster tank or the many fragments from the tank after it has been deliberately exploded, the very many pieces of chaff that can be launched in space to accompany the reentry vehicle, solid fuel fragments, separation debris, and the deliberate use of inflated decoys that resemble the reentry vehicle. One of the fundamental challenges in ballistic missile defense, therefore, is to determine which element in the multielement threat complex is the lethal object to be destroyed or made ineffective.[143] If the reentry vehicle is engaged by the defense after it reenters the earth's atmosphere, the drag introduced by the atmosphere causes the lightweight penetration aids to slow down much faster than the heavier high-speed reentry body. Thus the atmosphere separates the warhead from the penetration aids. A serious limitation of this tactic is that it results in a relatively small defended area. If, on the other hand, the ballistic missile threat is engaged outside of the atmosphere so as to achieve a large defended area, then there needs to be some way in which the threatening warhead can be distinguished from the "junk" that accompanies it. The use of radar in such a role is called *ballistic missile target discrimination.*

Meteorological Observation The original weather radars used for many years by the National Weather Service determined the presence of rainfall and estimated the rainfall rate. These were replaced in the early 1990s with the *S*-band Nexrad (WSR-88D) doppler weather radar,[144] which was designed to obtain much more information about the weather, other than it is raining somewhere. In addition to determining rainfall as a function of azimuth and height, Nexrad employs the doppler frequency shift to estimate wind speed as a function of direction and height in order to detect and measure damaging winds, severe turbulence, dangerous wind shear, and recognize the onset of tornadoes. Nexrad

differentiates hail from heavy rainfall, determines the tops of thunderstorms (an indicator of the intensity of such storms), and detects and tracks severe storms and mesocyclones. There are about 40 different weather products that can be generated by the WSR-88D radar.[145] These are processed automatically and displayed to an operator for further action. The extensive processing is such that the trained operator need not be a professional meteorologist to readily employ the information presented by the radar. Nexrad is a good example that radar need no longer be a "blob" detector!

Battlefield Surveillance The *X*-band Joint STARS airborne radar system employs a SAR mode for generating a maplike image of the battlefield and a GMTI mode for detecting moving targets. These two modes allow the recognition of terrain features, roads, fixed structures, military forces, artillery, and stationary as well as moving vehicles. Detection of moving vehicles is readily accomplished with modern GMTI methods. The radar detection of stationary ground targets of military interest is a classical radar application, but recognizing what type of target is being seen by the radar can be difficult. To be useful for target recognition, a radar should be able to distinguish tanks from trucks and from artillery. It would be even better if the type of tank or type of artillery could be determined.

Vehicles are difficult to recognize with radar since they have only a relatively few distinctive scattering centers. This is in contrast to a ship target that has many scattering centers, which makes recognizing one class of ship from many other ship classes achievable based on its ISAR image. Although the writer is not aware of a documented source defining the minimum number of distinctive scatterers needed for classifying targets, it appears that perhaps a dozen (more or less) individual scatterers might be necessary in a target image for reliable classification. The minimum number will depend, of course, on the type of target and the number of different target classes that have to be sorted. The number of scatterers defining a target might have to be greater if there are a large number of different classes of targets that have to be distinguished from one another. With both military aircraft and ships, the number of different classes that might be encountered range from many tens of target classes to more than a hundred if worldwide operation is expected.

As mentioned previously, determining the type of military land target seen by a SAR is called *automatic target recognition,* or ATR.[146] Radar ATR of land targets of military interest is complicated by the military tactic of hiding targets among the trees and other vegetation that can contaminate the target's echo signature as well as interfere with detection when the clutter echo is large. Thus ATR involves some method of STI (stationary target indication) to detect targets in the midst of clutter as well as foliage penetration. Foliage penetration requires operation at the lower radar frequencies, such as VHF, where the attenuation in propagating through trees and underbrush is lower than at microwave frequencies. Recognition of targets also requires very wide bandwidths to isolate the target's scatterers. Wide bandwidth combined with VHF results in ultrawideband (UWB) SAR.[147]

The interpretation of SAR images of terrain and populated areas has also been used in the civilian world for adding to the base of Geographic Information Systems (GIS), which use such information for land usage and planning, urban development, natural resource management, and topography.

Other Applications of Radar Information Extraction Radar has been used for recognizing one species of birds from another by means of the modulation on the bird's radar echo signals thought to be due to the characteristic beating of its wings.[148] Doppler radar can be used to detect the heartbeat and breathing of humans, even when located behind a cinder-block wall.[149,150] Radars can readily detect buried land mines, but the challenge has been to recognize the echo of a mine from the echoes from the many other underground objects that can be present, such as rocks, tree roots, debris, and changes in the underground surface characteristics.

REFERENCES

1. Skolnik, M. 1. "Radar Information from the Partial Derivatives of the Echo Signal Phase from a Point Scatterer." Naval Research Laboratory, Washington, D.C., Memorandum Rep. 6148, February 17, 1988.
2. Evans, J. V., and T. Hagfors. *Radar Astronomy,* New York: McGraw-Hill, 1968, Sec. 5.9.
3. Jurgens, R. F. "Earth-Based Radar Studies of Planetary Surfaces and Atmospheres." *IEEE Trans.* GE-20 (July 1982), pp. 293–305.
4. McDonough, R. N., and A. D. Whalen. *Detection of Signals in Noise.* San Diego, CA: Academic, 1995.
5. Goldman, S. *Frequency Analysis, Modulation, and Noise.* New York: McGraw-Hill, 1948, p. 281.
6. Slepian, D. "Estimation of Signal Parameters in the Presence of Noise." *IRE Trans.* no. PGIT-3 (March 1954), pp. 68–89.
7. Manasse, R. "Range and Velocity Accuracy from Radar Measurements," unpublished internal report dated February 1955, MIT Lincoln Laboratory, Lexington, MA. (Not generally available.)
8. Mallinckrodt, A. J., and T. E. Sollenberger. "Optimum Pulse-Time Determination." *IRE Trans.* no. PGIT-3 (March 1954), pp. 151–159.
9. Weiss, A. J. "Composite Bound on Arrival Time Estimation Errors." *IEEE Trans.* AES-22 (November 1986), pp. 751–756.
10. Torrieri, D. J. "Arrival Time Estimation by Adaptive Thresholding." *IEEE Trans.* AES-10 (March 1974), pp. 178–184.
11. Torrieri, D. J. "Adaptive Thresholding Systems." *IEEE Trans.* AES-13 (May 1977), pp. 273–280.
12. Ho, K. C., Y. T. Chan, and R. J. Inkol. "Pulse Arrival Time Estimation Based on Pulse Sample Ratios." *IEE Proc.* 142, Pt. F (August 1995), pp. 153–157.
13. Varian, R. H., W. W. Hansen, and J. R. Woodyard. "Object Detecting and Locating System." U.S. Patent 2,435,615, February 10, 1948.

14. Skolnik, M. *Introduction to Radar Systems,* 2nd ed. New York: McGraw-Hill, 1980, Sec. 3.5.

15. Wadley, T. L. "Electronic Principles of the Tellurometer." *Trans. South African Inst. Elec. Engrs.* 49 (May 1958), pp. 143–161; discussion pp. 161–172.

16. Manasse, R. "Range and Velocity Accuracy from Radar Measurements." Internal memo dated February 1955, MIT Lincoln Laboratory, Lexington, MA. (Not generally available.)

17. *Encyclopaedia Britannica* vol. 12, 15th ed., Chicago, IL, 1991, p. 125.

18. Sinsky, A. 1., and C. P. Wang. "Standardization of the Definition of the Radar Ambiguity Function." *IEEE Trans.* AES-10 (July 1974), pp. 532–533.

19. IEEE Standard 686-1990, Radar Definitions, April 20, 1990.

20. Woodward, P. M. *Probability and Information Theory, with Applications to Radar.* New York: McGraw-Hill, 1953, chap. 7.

21. Siebert, W. McC. "A Radar Design Philosophy." *IRE Trans.* IT-2 (September 1956), pp. 204–221.

22. Deley, G. W. "Waveform Design." In *Radar Handbook,* 1st ed., M. Skolnik (Ed.). New York: McGraw-Hill, 1970, p. 3-17.

23. Rihcazek, A. W. *Principles of High-Resolution Radar.* New York: McGraw-Hill, 1969, Sec. 5.3.

24. Woodward, P. M. Ref. 20, Sec. 7.2.

25. Dicke, R. H. "Object Detection System." U.S. patent no. 2,624,876, issued Jan. 6, 1953.

26. Klauder, J. R., et al. "The Theory and Design of Chirp Radars." *Bell System Technical J.* 39 (July 1960), pp. 745–808.

27. Cook, C. E., and M. Bernfeld. *Radar Signals.* Norwood, MA: Artech House, 1993.

28. Skolnik, M. *Introduction to Radar Systems.* 2nd ed. New York: McGraw-Hill, 1980, Sec. 3.3.

29. Borkowski, T. T. "Solid-State Transmitters." In *Radar Handbook,* 2nd ed. M. Skolnik (Ed.) New York: McGraw-Hill, 1990, Chap. 5.

30. Cook, C. E., and M. Bernfeld. Ref. 27, Sec. 7.3.

31. Caputi, W. J., Jr. "Stretch: A Time-Transformation Technique." *IEEE Trans.* AES-7 (March 1971), pp. 269–278.

32. Holt, D. J., and M. B. Fishwick. "Analog Waveform Generation and Processing." *Electronic Progress* 17, no. 1 (Spring 1975), pp. 2–16. (Published by Raytheon, Lexington, MA.)

33. Brookner, E. *Aspects of Modern Radar.* Norwood, MA: Artech House, 1988, pp. 25–28.

34. Farnett, E. C., T. B. Howard, and G. H. Stevens. "Pulse-Compression Radar." In *Radar Handbook,* 1st ed., M. Skolnik (Ed.). New York: McGraw-Hill, 1970, chap. 20.

35. Hartmann, C. S. "SAW Device Technology: Recent Advances and Future Trends." *Microwave J.* 1990 State of the Art Reference, pp. 73–89.

36. Maines, J. D., and E. G. S. Paige. "Surface-Acoustic-Wave Devices for Signal Processing Applications." *Proc. IEEE* 64 (May 1976), pp. 639–652.

37. Cambell, Colin. *Surface Acoustic Wave Devices and Their Signal Processing Applications.* New York: Academic, 1989.

38. Information obtained from Anderson Laboratories, Bloomfield, CT, brochure titled "Dispersive Delay Lines," 1986.

39. Martin, T. A. "Low Sidelobe IMCON Pulse Compression." *Proc. 1976 IEEE Ultrasonics Symposium,* pp. 411–414, IEEE Cat. #76 CH1120-5SU.

40. Farnett, E. C., and G. H. Stevens. "Pulse Compression Radar." In *Radar Handbook,* 2nd ed. M. Skolnik (Ed.). New York: McGraw-Hill, 1990, chap. 10.

41. Wehner, D. R. *High-Resolution Radar,* 2nd ed. Boston, MA: Artech House, 1995, Sec. 4.7.

42. MacWilliams, F. J., and N. J. A. Sloan. "Pseudo-Random Sequences and Arrays." *Proc. IEEE* 64 (December 1976), pp. 1715–1729.

43. Farnett, E. C., and G. H. Stevens. Ref. 40, Sec. 10.6.

44. Golomb, S. W. *Shift Register Sequences.* rev. ed. Laguna Hills, CA: Aegean Park, 1982.

45. Taylor, S. A., and J. L. MacArthur. "Digital Pulse Compression Radar Receiver." *APL Technical Digest* 6 (March/April 1967), pp. 2–10.

46. Turin, R. "Sequences with Small Correlation." In *Error Correcting Codes.* H. B. Mann (Ed.). New York: John Wiley, 1968, pp. 195–228.

47. Linder, J. "Binary Sequences Up to Length 40 With Best Possible Autocorrelation Function." *Electron. Lett.* 11 (October 10, 1975), p. 507.

48. Kerdock, A. M., R. Mayer, and D. Bass. "Longest Binary Pulse Compression Codes with Given Peak Sidelobe Levels." *Proc. IEEE* 74 (February 1986) p. 366.

49. Taylor, J. W., Jr., and H. J. Blinchikoff. "Quadriphase Code—A Radar Pulse Compression Signal with Unique Characteristics." *IEEE Trans.* AES-24 (March 1988), pp. 156–170.

50. Levanon, N., and A. Freedman. "Ambiguity Function of Quadriphase Coded Radar Pulse." *IEEE Trans.* AES-25 (November 1989), pp. 848–853.

51. Frank, F. L. "Polyphase Codes with Good Nonperiodic Correlation Properties." *IEEE Trans.* IT-9 (January 1963), pp. 43–45.

52. Cook and Bernfeld, Ref. 27, Sec. 8.4.

53. Nathanson, F. E. *Radar Design Principles.* 2nd ed. New York: McGraw-Hill, 1991, Sec. 12.5.

54. Lewis, B. L., F. F. Kretschmer, Jr., and W. W. Shelton. *Aspects of Radar Signal Processing.* Norwood, MA: Artech House, 1986, Chap. 2.

55. Lewis, B. L. "Range-Time Sidelobe Reduction Technique for FM-Derived Polyphase PC Codes." *IEEE Trans.* AES-29 (July 1993), pp. 834–840.

56. Nathanson, F. E. Ref. 53, Sec. 13.5.

57. Costas, J. P. "A Study of a Class of Detection Waveforms Having Nearly Ideal Range-Doppler Ambiguity Properties." *Proc. IEEE* 72 (August 1984), pp. 996–1009.

58. Chang, W., and K. Scarbrough. "Costas Arrays with Small Number of Cross-Coincidences." *IEEE Trans.* AES-25 (January 1989), pp. 109–112.

59. Golomb, S. W., and H. Taylor. "Constructions and Properties of Costas Arrays." *Proc. IEEE* 72 (September 1984), pp. 1143–1163.

60. Nathanson, F. E. Ref. 53, Sec. 13.11.

61. Farnett, E. C., and G. H. Stevens. Ref. 40, Sec. 10.4.

62. Cook, C. E., and M. Bernfeld. Ref. 27, Sec. 6.8.

63. Minkoff, J. *Signals, Noise, and Active Sensors.* New York: John Wiley, 1992, Sec. 9.2.

64. Thor, R. C. "A Large Time-Bandwidth Pulse-Compression Technique." *IEEE Trans.* MIL-6 (April 1962), pp. 169–173. Reprinted in *Radars* vol. 5, Pulse Compression, D. K. Barton (Ed.). Boston, MA: Artech House, 1875.

65. Kroszcynski, J. J. "Pulse-Compression by Means of Linear-Period Modulation." *Proc. IEEE* 57 (July 1969), pp. 1260–1266.

66. Minkoff, J. Ref. 63, Sec. 9.3.

67. Rowlands, R. O. "Detection of a Doppler-Invariant FM Signal by Means of a Tapped Delay Line." *J. Acoust. Soc. Am.* 37 (April 1965), pp. 608–615.

68. Altes, R. A., and E. L. Titlebaum. "Bat Signals as Optimally Doppler Tolerant Waveforms." *J. Acoust. Soc. Am.* 48 (1970), pp. 1014–1020.

69. Altes, R. A. "Radar/Sonar Acceleration Estimation with Linear-Period Modulated Waveforms." *IEEE Trans.* AES-26 (November 1990), pp. 914–923.

70. Kramer, S. A. "Doppler and Acceleration Tolerances of High-Gain, Wideband Linear FM Correlation Sonars." *Proc. IEEE* 55 (May 1967), pp. 627–636.

71. Golomb, S. W. Ref. 44, Chap. VI.

72. Belyayev, V. S. "A New Pseudorandom, Phase-Controlled Signal Based Upon a Nonlinear Sequence and the Possibilities for Generating It." *Radiophysics and Quantum Electronics* 34 (March 1991), pp. 285–287.

73. Golay, M. J. E. "Complementary Series." *IRE Trans.* IT-7 (June 1960), pp. 82–87.

74. Levanon, N. *Radar Principles.* New York: John Wiley, 1988, pp. 159–162.

75. Cloke, J. A. "Ambiguity Function of Complementary Series." *IEE International Conf., Radar-82,* October 18–20, 1982, pp. 477–481.

76. Welti, G. R. "Quaternary Codes for Pulse Radar." *IRE Trans.* IT-7 (June 1960), pp. 400–408.

77. Nathanson, F. E. Ref. 53, p. 564.

78. Cook, C. E., and M. Bernfeld. Ref. 27, pp. 264–269.

79. Kretschmer, F. F., and F. C. Lin. "Huffman-Coded Pulse Compression Waveforms." Naval Research Laboratory, Washington, D.C., Report 8894, May 23, 1985.

80. Eves, J. L., and E. K. Reedy. *Principles of Modern Radar.* Van Nostrand Reinhold, New York, 1987, Sec. 15.3.2.

81. Key, E. L., E. N. Fowle, and R. D. Haggarty. "A Method of Sidelobe Suppression in Phase-Coded Pulse Compression Systems." MIT Lincoln Laboratory, Lexington, MA, TR-209, August 28, 1959.

82. Nathanson, F. E. Ref. 53, Sec. 12.4.

83. Ackroyd, M. H., and F. Ghani. "Optimum Mismatched Filters for Sidelobe Suppression." *IEEE Trans.* AES-9 (March 1973), pp. 214–218.

84. Eves and Reedy, Ref. 80, Sec. 15.5.2.

85. Morgan, G. B., P. Dassanayake, and O. A. Liberg. "The Design and Performance of Transversal Filters for Sidelobe Reduction of Pulses Compressed from Combined Barker Phase Codes." *The Radio and Electronic Engineer* 51 (June 1981), pp. 272–280.

86. Nathanson, F. E. Ref. 53, pp. 556–557.

87. Baden, J. M., and M. N. Cohen. "Optimal Peak Sidelobe Filters for Biphase Pulse Compression." *Record of the 1995 IEEE International Radar Conf.,* IEEE Catalog No. 90CH2882-9, pp. 249–252.

88. Castella, F. R., and S. A. Rudie. "Detection Performance of Phase-Coded Radar Waveforms with Various Types of Limiting." *IEE Proc.* 136, Pt. F (June 1989), pp. 118–121.

89. Cahn, C. R. "A Note on Signal-to-Noise Ratio in Band-Pass Limiters." *IEEE Trans.* IT-7 (January 1961), pp. 39–43.

90. Bogotch, S. E., and C. E. Cook. "The Effect of Limiting on the Detectability of Partially Time-Coincident Pulse Compression Signals." *IRE Trans.* MIL-9 (January 1965), pp. 17–24.

91. Woerrlein, H. H. "Capture and Spurious Target Generation Due to Hard Limiting in Large Time-Bandwidth Product Radars." Naval Research Laboratory, Washington, D.C., Report 7001, December 22, 1969.

92. Maric, S. V., I. Seskar, and E. L. Titlebaum. "On Cross-Ambiguity Properties of Welsh-Costas Arrays." *IEEE Trans.* AES-30 (October 1994), pp. 1063–1071.

93. Deng, H. "Synthesis of Binary Sequences with Good Autocorrelation and Cross-correlation Properties by Simulated Annealing." *IEEE Trans.* AES-32 (January 1996), pp. 98–107.

94. Shrader, W. W., and V. Gregers-Hansen. *Radar Handbook,* M. Skolnik (Ed.). New York: McGraw-Hill, 1990, "MTI Radar," Chap. 15, pp. 15.55–15.57.

95. Arthur, J. W. "SAW Pulse Compression in Modern Multi-Channel Radar Applications." *Microwave J.* 29 (January 1986), pp. 159–169.

96. Dixon, R. C. *Spread Spectrum Systems.* New York: Wiley Interscience, 1976.

97. Geffe, P. R. "Open Letter to Communications Engineers." *Proc. IEEE* 55 (December 1967), p. 2173.

98. Skolnik, M. I. "Radar's Environmental Role." *IEEE Potentials* 10 (April 1991), pp. 13–16.

99. Linde, G. J. "Use of Wideband Waveforms for Target Recognition with Surveillance Radars." *Record of the IEEE 2000 International Radar Conf.* May 7–12, 2000, Washington, D.C., pp. 128–133. See also Linde, G. J., and C. V. Platis. "Target Recognition with Surveillance Radar." *NRL Review.* Naval Research Laboratory, Washington, D.C., pp. 118–120, 1995.

100. Hudson, S., and D. Psaltis. "Correlation Filters for Aircraft Identification from Radar Range Profiles." *IEEE Trans.* AES-29 (July 1993), pp. 741–748.

101. Zyweck, A., and R. E. Bogner. "Radar Target Classification of Commercial Aircraft." *IEEE Trans.* AES-32 (April 1996), pp. 598–606.

102. Cutrona, L. J. "Synthetic Aperture Radar." *Radar Handbook,* M. Skolnik (Ed.). New York: McGraw-Hill, 1990, Chap. 21.

103. Kerr, D., S. Musman, and C. Bachmann. "Automatic Recognition of ISAR Ship Images." *IEEE Trans.* AES-32 (October 1996), pp. 1392–1404.

104. Jain, A., and I. Patel. "Dynamic Imaging and RCS Measurements of Aircraft." *IEEE Trans.* AES-31 (January 1995), pp. 211–226.

105. Harris, et al. "Dynamic Air-to-Air Imaging Measurement System." *Conf. Proceedings of the 14th Annual Meeting of the Antenna Measurements Techniques Association,* October 19–23, 1992, pp. 6-11 to 6-16.

106. Steinberg, B. D. "Microwave Imaging of Aircraft." *Proc. IEEE* 76 (December 1988), pp. 1578–1592.

107. Steinberg, B. D., D. L. Carlson, and W. Lee. "Experimental Localized Radar Cross Sections of Aircraft." *Proc. IEEE* 77 (May 1989), pp. 663–669.

108. Dinger, R., et al. "Measurements of the Radar Cross Section and Inverse Synthetic Aperture Radar (ISAR) Images of a Piper Navajo at 9.5 GHz and 49 GHz." Naval Command Control and Ocean Surveillance Center (NRaD), Tech. Rep. 1569, January 1993.

109. Pettengill, G. H. "Radar Astronomy." In *Radar Handbook,* 1st ed. M. Skolnik (Ed.). New York: McGraw-Hill, 1970, chap. 33.

110. Hynes, R., and R. E. Gardner. "Doppler Spectra of S-Band and X-Band Signals." Supplement to *IEEE Trans.* AES-3 (November 1967), pp. 356–365. Also, *Report of NRL Progress* (January 1968), pp. 1–10.

111. Dunn, J. H., and D. D. Howard. "Target Noise." *Radar Handbook,* 1st ed. M. Skolnik (Ed.). New York: McGraw-Hill, 1970, Chap. 28, Sec. 28.5.

112. Collot, G. "Fixed/Rotary Wings Classification/Recognition." *Proc. of the CIE International Conf. on Radar,* Beijing, China, October 22–24, 1991, pp. 610–612.

113. Bullard, B. D., and P. C. Dowdy. "Pulse Doppler Signature of a Rotary-Wing Aircraft." *Proc. 1991 IEEE National Radar Conf.,* Los Angeles, CA, March 12–13, 1991, pp. 160–163.

114. Fliss, G. G., and D. L. Mensa. "Instrumentation for RCS Measurements of Modulation Spectra of Aircraft Blades." *Proc. IEEE 1986 National Radar Conf.,* pp. 95–99.

115. Kulpa, K., Z. Czekala, J. Misiurewicz, and J. Falkiewicz. "Parametric Detection of the Helicopter Hub Echo." *Proc. 1999 IEEE Radar Conf.,* Waltham, MA, pp. 262–266, IEEE Catalog No. 99CH36249.

116. Gardner, R. E. "Doppler Spectral Characteristics of Aircraft Radar Targets at S-Band." Naval Reserach Laboratory, Washington, D.C., Report 5656, August 3, 1961.

117. Bell, M. R., and R. A. Grubbs. "JEM Modeling and Measurement for Radar Target Identification." *IEEE Trans.* AES-29 (January 1993), pp. 73–87.

118. Copeland, J. R. "Radar Target Classification by Polarization Properties." *Proc. IRE* 48 (July 1960), pp. 1290–1296.

119. Holm, W. A. "Polarimetric Fundamentals and Techniques." *Principles of Modern Radar,* J. L. Eaves and E. K. Reedy (Eds.). New York: Van Nostrand Reinhold, 1987, chap. 20.

120. Evans, D. L., T. G. Farr, J. J. Van Zyl, and H. A. Zebker. "Radar Polarimetry: Analysis Tools and Applications." *IEEE Trans.* GRS-26 (November 1988), pp. 774–789.

121. Boerner, W. M., W-L Yan, A-Q Xi, and Y. Yamaguchi. "On the Basic Principles of Radar Polarimetry: the Target Characteristic Polarization State Theory of Kennaugh, Huynen's Polarization Fork Concept, and Its Extension to the Partially Polarized Case." *Proc. IEEE* 79 (October 1991), pp. 1538–1550.

122. Giuli, D. "Polarization Diversity in Radars." *Proc. IEEE* 74 (February 1986), pp. 245–269.

123. Howard, D. D. "High Range-Resolution Monopulse Radar." *IEEE Trans.* AES-11 (September 1975), pp. 749–755.

124. Ksienski, A. A., Y. T. Lin, and L. J. White. "Low-Frequency Approach to Target Identification." *Proc. IEEE* 63 (December 1975), pp. 1651–1660.

125. Lin, H., and A. A. Ksienski. "Optimum Frequencies for Aircraft Classification." *IEEE Trans.* AES-17 (September 1981), pp. 656–665.

126. Chuang, C. W., and D. L. Moffatt. "Natural Resonances of Radar Targets via Prony's Method and Target Discrimination." *IEEE Trans.* AES-12 (November 1976), pp. 583–589.

127. Moffatt, D. L., and R. K. Mains. "Detection and Discrimination of Radar Targets." *IEEE Trans.* AP-23 (May 1975), pp. 358–367.

128. Baum, C. E., E. J. Rothwell, K-M Chen, and D. P. Nyquist. "The Singularity Expansion Method and Its Application to Target Identification." *Proc. IEEE* 79 (October 1991), pp. 1481–1492.

129. Dudley, D. G. "Progress in Identification of Electromagnetic Systems." *IEEE Ant. and Prop. Society Newsletter* (August 1988), pp. 5–11.

130. Fok, F. Y. S., and D. L. Moffatt. "The K-Pulse and E-Pulse." *IEEE Trans.* AP-35 (November 1987), pp. 1325–1326.

131. Chen, K-M, D. P. Nyquist, E. J. Rothwell, L. L. Webb, and B. Drachman. "Radar Target Discrimination by Convolution of Radar Return with Extinction-Pulse and Single-Mode Extraction Signals." *IEEE Trans.* AP-34 (July 1986), pp. 896–904.

132. Optiz, C. L. "Metal-Detecting Radar Rejects Clutter Naturally." *Microwaves* 15 (August 1976), pp. 12–14.

133. Harger, R. G. "Harmonic Radar Systems for Near-Ground In-Foliage Nonlinear Scatterers." *IEEE Trans.* AES-12 (March 1976), pp. 230–245.

134. Flemming, M. A., F. H. Mullins, and A. W. D. Watson. "Harmonic Radar Detection Systems." *IEE RADAR-77 International Conf.,* October 25–28, pp. 552–554, 1977.

135. Powers, E. J., J. Y. Hong, and Y. C. Kim. "Cross Sections and Radar Equation for Nonlinear Scatterers." *IEEE Trans.* AES-17 (July 1981), pp. 602–605.

136. Hong, J. Y., and E. J. Powers. "Detection of Weak Third Harmonic Backscatter from Nonlinear Metal Targets." *IEEE Eascon-83,* September 19–21, 1983, pp. 169–175, 83CH1967-9, ISSN:0531-6863.

137. Bahr, A. J., and J. P. Petro. "On the RF Frequency Dependence of the Scattered Spectral Energy Produced by Intermittent Contacts Among the Elements of a Target." *IEEE Trans.* AP-25 (July 1978), pp. 618–621.

138. Newburgh, R. G. "Basic Investigation of the RADAM Effect." Rome Air Development Center, Rome, N.Y., Report RADC-TR-151, June 1978.

139. Trunk, G. V., and J. D. Wilson. "Association of DF Bearing Measurements With Radar Tracks." *IEEE Trans.* AES-23 (July 1987), pp. 438–447.

140. Saha, R. K. "Analytical Evaluation of an ESM/Radar Track Association Algorithm." *SPIE* 1698, "Signal and Data Processing of Small Targets." (1992), pp. 338–347.

141. Farina, A., and B. La Scala. "Methods for the Association of Active and Passive Tracks for Airborne Sensors." *International Radar Symposium,* Munich, Germany, September, 1998.

142. Hawkins, C. F. "Friendly Fire: Facts, Myths and Misperceptions." *Proc. U.S. Naval Institute* 120 (June 1994), pp. 54–59.

143. Silberman, G. L. "Parametric Classification Techniques for Theater Ballistic Missile Defense." *Johns Hopkins APL Technical Digest* 19, no. 3 (1998), pp. 322–339.

144. Heiss, W. H., D. L. McGrew, and D. Sirmans. "Nexrad: Next Generation Weather Radar (WSR-88D)." *Microwave J.* 33 (January 1990), pp. 79–98.

145. Crum, T. D., and R. L. Alberty. "The WSR-88D and the WSR-88D Operational Support Facility." *Bulletin of the American Meteorological Society* 74 (September 1993), pp. 1669–1687.

146. Dudgen, D. E., and R. T. Lacoss. "An Overview of Automatic Target Recognition." MIT Lincoln Laboratory Journal, "Special Issue on Automatic Target Recognition," vol. 6, pp. 3–10, Spring 1993.

147. Sheen, D. R. et al. "The P-3 Ultra-Wideband SAR: Description and Examples." *Proc. 1996 IEEE National Radar Conf.*, pp. 50–53, IEEE Catalog no. 96CH35891.

148. Vaughn, C. R. "Birds and Insects as Radar Targets: A Review." *Proc. IEEE* 73 (February 1985), pp. 205–227.

149. Chen, K-M, et al. "An X-Band Microwave Life-Detection System." *IEEE Trans.* BME-33 (July 1986), pp. 697–701.

150. Geisheimer, J. "A Radar System for Monitoring Human Vital Signs." *IEEE Potentials* 17 (January 1999), pp. 21–24.

PROBLEMS

6.1 (a) Sketch the rms range error (in meters) for a quasi-rectangular pulse with a half-power pulse width of 2 μs, as a function of the peak-signal-to-mean-noise ratio ($2E/N_0$) over the range of values from 10 to 60 dB. (b) Why might it not be appropriate to consider signal-to-noise ratios below 10 dB and above 60 dB?

6.2 Derive the rms error in measuring the time delay for a gaussian pulse of half-power width τ [Eq. (6.17)].

6.3 Based on the measurement of doppler frequency shift, sketch the rms error of the radial velocity (in meters) as a function of the width τ of a rectangular pulse when the pulse width varies from 1 μs to 10 ms for (a) constant pulse energy and (b) constant peak power. The frequency is 5400 MHz. Assume in both cases that $2E/N_0 = 36$ when the pulse width is 1 μs.

6.4 (a) What is the minimum width τ of a rectangular pulse that can be used with an X-band radar (9375 MHz) if it is desired to achieve a 10 kt radial velocity accuracy (based on the doppler frequency measured by a single pulse), when $2E/N_0 = 23$ dB? (b) What is the minimum range (in nautical miles) that corresponds to this pulse width? (c) In part (a) of this question, what should be the value of $2E/N_0$ (in dB) to achieve a 10 kt radial velocity accuracy if the pulse width can be no longer than 10 μs? (d) What would be the minimum pulse width in (a) if the radar operated at W band (94 GHz)? (e) Comment on the utility of accurately measuring the velocity with a single short pulse.

6.5 There are two methods for finding the radial velocity of a target. One is based on the doppler shift $f_d = 2v_r/\lambda$; the other is based on the rate of change of range with time $\Delta R/\Delta t$. They give different measurement accuracies. (a) What is the expression for the radial velocity error, δv_d, found by measuring the doppler frequency shift of a long quasi-rectangular pulse of width τ and RF frequency f_0? (b) What is the expression for the radial velocity error, δv_r, found from the rate of change of range based on two range measurements R_1 and R_2 separated by a time τ, so that the velocity is $v_r = (R_2 - R_1)/\tau$, and τ is

the same as the pulse width of the doppler measurements? The pulses in this range-rate measurement are of gaussian shape with a half-power bandwidth B [use right-hand side of Eq. (6.17)]. Assume the total value of $2E/N_0$ in each of these two methods for radial velocity measurement [(a) and (b)] is the same. (c) What is the value of $\delta v_d/\delta v_r$? (d) Based on your answer in (c), which is the more accurate method of velocity measurement, the doppler method or the range-rate method? (e) Under what conditions will these two methods give comparable accuracies (assuming the same total $2E/N_0$)? (f) Why do you think the doppler method has not been used very often for a velocity measurement?

6.6 (a) What value of $2E/N_0$ (in dB) is required to achieve an angular accuracy of 0.3 mrad when the antenna beamwidth is one deg, assuming the antenna has a cosine aperture illumination? (b) If the signal received by this radar antenna has a value of $2E/N_0 = 23$ dB for a particular target at a range of 150 nmi, at what range will the radar first be capable of obtaining an angular accuracy of 0.3 mrad?

6.7 Determine $\beta\alpha$ (the product of the effective pulse width β and the effective time duration α) for the following: (a) gaussian pulse, (b) triangular pulse, and (c) quasi-rectangular pulse. (Use the expressions for β and α already given in the text.) (d) Based on the above answers, how much flexibility is there in selecting one of these simple waveforms to achieve a large value of $\beta\alpha$ for the purpose of making an accurate measurement of both time delay and frequency? (e) What option, other than a large $\beta\alpha$, is available for obtaining an accurate measurement of both time delay and frequency simultaneously?

6.8 Derive the rms error in measuring frequency for a rectangular pulse of width τ (Eq. 6.24).

6.9 How is the rms error in measuring frequency affected when the width of a rectangular pulse is increased by a factor of four, with the peak power remaining constant?

6.10 (a) A K-band (24.15 GHz) radar speed-gun, used for measuring the speed of an auto or a baseball, has a claimed accuracy of 0.1 statute mile/hour. If the signal-to-noise ratio ($2E/N_0$) is 17 dB, what observation time is needed to achieve this accuracy? (b) How far does a car with a speed of 60 mph travel during this time? (c) Assuming the same obervation time and value of $2E/N_0$ as in part (a), what would be the accuracy of an X-band (10.525 GHz) radar speed-gun?

6.11 Using the first part of Eq. 6.21, $\beta^2 E = \int_{-\infty}^{\infty} [s'(t)]^2\, dt$, show that the rms error in measuring the time delay of a trapezoidal pulse is the same as that given by Eq. (6.14).

6.12 Note that the sine integral function $Si(4\pi) = 1.492$. (The energetic reader might try to find the time waveform $s(t)$ out of a rectangular filter when $B_s\tau_r = 4$, plot $s(t)$ to find its half-power width τ, and express β and α in terms of the half-power width τ.) Find the values of the effective bandwidth β and the effective time duration α for a different quasi-rectangular pulse formed by taking the bandwidth B_s [in Fig. 6.2, Eq. (6.9), and Eq (6.25)] extending from $-2/\tau_r$ to $+2/\tau_r$, where τ_r is the width of the original rectangular pulse. Thus $B_s\tau_r = 4$.

6.13 Using the definition of β [Eq. (6.8)], definition of α [Eq. (6.23)], Eq. (6.21), Parseval's relation [Eq. (5.13)], Schwartz inequality, [Eq. (5.11)], and performing integration by parts, show that $\beta\alpha \geq \pi$ and that the equality holds for the gaussian function. (Only for those who enjoy a mathematical exercise.)

6.14 Show that $\int (2\pi f)^2 |S(f)|^2\, df = -\int s''(t)s(t)dt = \int [s'(t)]^2\, dt$, which is Eq. (6.21). The limits of all integrals are from $-\infty$ to $+\infty$. (Start with $s(t)$ expressed as an inverse Fourier transform of $S(f)$,and differentiate $s(t)$ twice.)

6.15 What is the message of the radar "uncertainty principle"?

6.16 Show that the rms error for the measurement of the phase of a sinewave is $\delta\phi = (2S/N)^{-1/2}$, where S/N is the signal-to-noise ratio. [This can be obtained from the error in making a measurement of the time when a sinewave crosses the time axis. Note that phase ϕ is equal to $2\pi ft$, so that $\delta\phi = 2\pi f\, \delta t$. The derivation is similar to that given for the rectangular pulse that led to Eq. (6.4), except that the time is measured when the sinewave crosses the time axis once.]

6.17 What information about the radar waveform can be obtained from the ambiguity diagram?

6.18 The amplitude of the peak of the signal out of a matched filter is $2E$, where E is the signal energy. This can be seen from Eq. (6.42). A peak output of $2E$ implies that the output has units of energy [(volts)^2 $\times$ time], but this cannot be so since the output of a filter should have volts as the unit. Show that the unit for the output of the matched filter is actually volts and not energy. (This is related to problem 5.6.)

6.19 Explain qualitatively (no derivation needed) why the ambiguity diagram shown in Fig. 6.9 is that of a down-chirp; i. e., frequency of the linear FM waveform decreases with time.

6.20 (a) Why might one want to use both an up-chirp and a down-chirp waveform on two successive pulses in a linear-FM pulse-compression radar system? (b) When might one use a waveform with the following three contiguous parts: (1) an unmodulated CW waveform, (2) a down-chirp and (3) an up-chirp? (c) Why might one want a radar to operate with an up-chirp when another similar nearby radar operates with a down-chirp?

6.21 A C-band (5.5 GHz) ballistic-missile detection radar employs a linear-FM pulse-compression radar with a 1.0-ms uncompressed pulse-width having a down-chirp covering a bandwidth of 200 MHz. If the target has a radial velocity of 2 km/s, what is the error in range (km) due to the doppler shift of the target? What is the range error in terms of the resolution of the compressed pulse?

6.22 Determine the shift in the time delay ΔT_r with a linear-FM pulse-compression waveform whose bandwidth is B and time duration is T when the target experiences a doppler frequency shift f_d for the following two cases: (a) Ballistic-missile-detection radar with $B = 1$ MHz and $T = 1$ ms, when $f_d = 100$ kHz. (b) Aircraft-detection radar with $B = 100$ MHz and $T = 10$ μs, when $f_d = 1$ kHz. In each of the two cases: (c) What is the doppler-caused time-shift error, measured in meters of range? (d) What is the ratio of the doppler-caused time shift and the resolution in time delay of the waveform (which you may assume to be $1/B$)?

6.23 How can one find the true range and true doppler frequency shift of a target when using linear-FM pulse-compression waveforms?

6.24 In this problem two nearby aircraft are both located within the antenna beam, but one is trailing slightly behind the other. The trailing aircraft has a slower speed than the leading

aircraft. They are being detected by a linear-FM pulse-compression radar. Describe why, in this case, it is better to use an up-chirp rather than a down-chirp when it is desired to provide a wide separation between the two output echo signals.

6.25 (a) Show that the tapped delay line that generates a Barker code of length 5 can be used as a matched filter if the receive signal is inserted from the opposite end. (This can be shown by sketching the output of the matched filter.) (b) How can the dispersive filter used to generate a linear-FM (chirp) signal be used as the matched filter to receive the echo signal?

6.26 Show (simple diagrams are okay) that the two different Barker codes of length 4 in Table 6.4 are also complementary codes.

6.27 *Background:* Generally, one can assume that the doppler shift is constant across the spectrum of the echo signal. There are cases, however, where this is not so, as in the current problem. *Problem:* A ballistic missile detection radar is attempting to detect a target moving with a radial velocity of 2 nmi/s. The radar employs a linear-FM pulse-compression waveform with a bandwidth of 100 MHz. What is the longest pulse width that can be used before significant degradation of the compressed pulse width occurs? (See the discussion in the subsection *Doppler-Tolerant Pulse-Compression Waveforms.*)

6.28 In the previous problem, one might have used the criterion that $2BTv_r/c$ must not exceed unity. This criterion is based on the doppler at the leading edge of a long FM pulse of length T and bandwidth B being significantly different from the doppler at the trailing edge of the pulse when the target radial velocity is v_r. Derive the expression $2BTv_r/c < 1$. [One criterion that might be applied is to require the spread in the time-delay measurements caused by the difference in the doppler shifts at the leading and trailing edges of the pulse of duration T to be no greater than the spread ($1/B$) in the time-delay measurement when the target is stationary (no doppler shift).]

6.29 A linear-FM pulse-compression radar has a bandwidth B and an uncompressed pulse duration T. The matched filter is followed by a sidelobe-reduction filter whose weighting function is

$$W(f) = \cos(\pi f/B)\ \text{rect}\ (f/B)$$

where rect $(x) = 1$ for $|x| < 1/2$ and equals 0 for $|x| > 1/2$. The signal spectrum at the output of the matched filter is rectangular and is given as

$$S_m(f) = \sqrt{T/B}\ \text{rect}\ (f/B)$$

and the noise spectrum at the matched filter output is

$$N(f) = (N_0/2)\ \text{rect}\ (f/B)$$

Find:

a. The signal waveform $s_w(t)$ at the output of the sidelobe-reduction filter.
b. The loss in signal-to-noise ratio due to the sidelobe-reduction filter.
c. The level of the first sidelobe of the output waveform $s_w(t)$.

(Acknowledgment: This problem was generously given to the author by an instructor of a radar course in Huntsville, Alabama many years ago. He used it as part of a take-home exam. Unfortunately, over the years I lost his name, but wish to acknowledge, with gratitude, my use of his problem here and throughout the years in my own radar course.)

6.30 An X-band (9.5 GHz) ground-based ISAR (inverse synthetic aperture radar) is imaging an aircraft at a range of 20 nmi. The aircraft is traveling at a speed of 250 kt and is on a tangential trajectory (i.e., it is perpendicular to the radar line of sight). The ISAR image is obtained as the aircraft is viewed broadside to the radar line of sight, so that the viewing aspect is 90° or in close vicinity to it. What must the total radar observation time be in order to obtain an image with a cross-range resolution of 1 m?

6.31 What effect might the echo from a rapidly rotating antenna located within a range-resolution cell of an ISAR have on the ISAR image of the target?

6.32 A ground-based S-band (3.2 GHz) radar is observing a helicopter. One may assume that the velocity of the tip of the blade of a helicopter is 210 m/s and that the length of a blade is 6 m.

a. What is the time between blade flashes for a two-blade and a three-blade helicopter?
b. What is the time duration of the blade flash for a two-blade and a three-blade helicopter?
c. If the radar antenna has an azimuth beamwidth of 3°, at what rpm must the antenna be rotated in order to insure that the blade flash will be detected on each scan of the antenna?
d. What must the prf of the radar be to insure that there are at least five pulses received from the blade flash?
e. What is the radar cross section of the blade flash for an odd number of blades?
f. If the helicopter is hovering, what is the maximum doppler shift that might be obtained from the helicopter blades?
g. What is the maximum forward speed (kt) that a helicopter can have if the actual velocity of the tip is to be less than 0.8 of Mach 1? (Mach 1 can be assumed to be 343 m/s.) What is the maximum doppler shift of the radar echo from the helicopter blade?

6.33 (a) What options are available for reliably recognizing one aircraft target from another by noncooperative methods? (b) Describe how one might reliably recognize a ship target by noncooperative methods. (c) Describe how one might reliably recognize a helicopter target by noncooperative methods.

6.34 It has sometimes been said that two pulse signals, one at frequency f_1 and the other at frequency f_2, each of width τ, can be resolved (separated by filters) by a radar if the number of cycles within the pulse width τ at frequency f_1 differs by at least one from the number of cycles within the pulse width τ at frequency f_2. (a) Show that this criterion for resolution is equivalent to the more usual criterion that $|f_1 - f_2| \geq 1/\tau$. (b) A long-range UHF radar (440 MHz) for the detection of satellites is assumed to have a pulse width of

2 ms. What is the minimum difference in radial velocity of two targets that will allow them to be separated by doppler filtering?

6.35 A linear-FM waveform and a $(\sin \pi ft)/\pi ft$ waveform both have a uniform power spectrum. Does that mean they both can use the same pulse compression filter? (More than a simple yes or no answer would be nice.)

6.36 It was said in Sec. 6.5 in the discussion of linear-FM pulse-compression that the true range of a target can be found by using both an up-chirp and a down-chirp waveform, and taking the average of the two time delays. It is also possible to obtain the true doppler frequency shift using the same two measurements of time delay from the up-chirp and down-chirp waveforms. (This was Problem 6.23.) Derive an expression for the rms accuracy of the frequency measurement found from the two time delay measurements when the frequency extent of the linear FM is B and the time duration is T.